Originally published at:
Meyer, Friederike. Neuroeconomic approaches to emotion-related influences on decision-making. 2015, University of Zurich, Wirtschaftswissenschaftliche Fakultät.

Neuroeconomic approaches to emotion-related influences on decision-making

Dissertation

submitted to the Faculty of Economics,

Business Administration and Information Technology

of the University of Zurich

to obtain the degree of
Doktorin der Neuroökonomie, Dr. sc.
(corresponds to Doctor of Neuroeconomics, PhD)

presented by

Friederike Meyer
from Germany

approved in July 2015 at the request of
Prof. Dr. Christian C. Ruff
Prof. Dr. Philippe N. Tobler

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 15/07/2015

Chairman of the Doctoral Board: Prof. Dr. Todd A. Hare

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my supervisor Christian Ruff for his support of my PhD studies and research, for his patience, motivation, and knowledge. Besides my supervisor, I would like to thank Philippe Tobler for investing time and effort in being the co-supervisor of this dissertation.

I would like to give special thanks to Tony Williams who patiently answered my questions concerning economics and statistics. He helped me out many times, without getting tired of giving repeated explanations.

Thomas Epper also supported me with data analysis and a critical view on data and design, always striving for stellar significance. Thank you!

Working life in BLU-104 would not have been as inspirational and entertaining without my officemates. Yosuke "Moritschi" Morishima, my sincerest gratitude for your scientific advice, personal support, and the fun times we spent together! Chaohui Guo, many thanks go to you four our helpful and motivating discussions, shared laughs and complaints!

Special thanks go to Azade Seid-Fatemi for our "general support system", including inspirational coffee breaks, file sharing, and crisis intervention. No distance is too short for a little Skype conversation!

Sonja Vogt, thanks for your honesty and little pushes in just the right moments! May I pay you back and push you towards a trip to Eastern Germany?

The studies presented in this dissertation were supported by the motivation and energy of many BLU and SNS-Lab colleagues. I would like to give many thanks to Marius Moisa, Daniela Pfabigan, Sophia Wagner, and Jan Engelmann for our collaborative work.

Nina Steinmann earns my special thanks for her unbelievably patient "dpi" support and friendship.

I would like to thank Julia Brinkkötter for spending way more time than expected proofreading this dissertation.

Finally, I would like to thank my family and friends for their everlasting support. My mother Annegret Meyer has shown incredible strength in supporting her children in every imaginable way.

Summary

Decades of classic economic research have neglected the role of incidental and integral emotional factors in human decision-making. Standard economic models assume that decision-making is consequentialist in nature: Decision-making is postulated to be guided by the decision maker's rational assessment of desirability and likelihood of alternative outcomes, i.e., by his strive to maximize utility (Rick & Loewenstein, 2008). However, advances in psychology and behavioral economics led to the gradual acceptance of incidental and integral emotion-related forces on decision-making. Still, we particularly lack detailed insight into the neurobiological mechanisms of the interaction between emotion and decision-making.

The purpose of this dissertation was to gain a more detailed understanding of these neurobiological substrates of the modulatory effects of different emotion-related factors on decision-making. To this end, four studies investigated different aspects of this interaction with behavioral, neuroimaging and neurostimulation tools.

Study A employed transcranial magnetic stimulation in order to investigate the neurobiological mechanisms causally underlying the framing effect, a decision-making bias hypothesized to be based on the influence of integral emotion on the decision-making process. This study shows that temporal disruption of activity in unilateral dorsolateral prefrontal cortex, a brain region known for its involvement in cognitive control, does not lead to changes in framing susceptibility, and casts doubt on the essential role of this brain region in the framing effect.

In Study B, the effects of incidental anticipatory anxiety on risky decision-making in a non-social context were investigated. The results demonstrate that induced anxiety leads to significant changes in neural value coding (especially in insula, ventral striatum and ventromedial prefrontal cortex), without changes in observable behavior (i.e., choices).

Study C focused on the effects of incidental anticipatory anxiety on social decision-making. Anticipatory anxiety caused a breakdown in trusting behavior, which was correlated with activity and connectivity changes in neural social cognition networks (e.g., temporoparietal junction, posterior superior temporal sulcus). The results of Studies B and C highlight the task- and context-specific effects of anxiety on behavioral and neural aspects of non-social and social risky decision-making.

Study D addressed the neuroanatomical substrate of emotion-related traits (i.e., personality and emotion-related traits) consistently found to influence daily decision-making. I found unique and specific associations between distinct affective and personality features such as impulsivity, sensation seeking, and negative emotionality and brain morphometry. Importantly, the identified brain regions have been shown to be (functionally and structurally) involved in decision-making and valuation. This finding emphasizes the association between affective and personality traits and variation in individual decision-making styles.

In sum, this dissertation contributes to a detailed understanding of the modulatory influence of emotion on decision-making. I present different neurofunctional and neurostructural correlates of emotion-decision-making interactions, which are highly task-dependent. Notably, these findings might offer a neurobiological account for deviant decision-making observed in numerous psychiatric disorders such as anxiety and affective disorders.

List of manuscripts

This dissertation is based on the following research articles:

Meyer, F., Pfabigan, D.M., Ruff, C.C. Framing effects on risky decision-making: Investigating a causal role of dorsolateral prefrontal cortex with TMS. *In preparation.*

Engelmann, J.B.*, Meyer, F.*, Fehr, E.†, Ruff, C.C.† (2015). Anticipatory anxiety disrupts neural valuation during risky choice. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 35*(7), 3085–3099. doi:10.1523/JNEUROSCI.2880-14.2015

*****shared first authorship, †shared last authorship

Engelmann, J.B., Meyer, F., Ruff, C.C.†, Fehr, E.† The neural circuitry of emotion-induced distortions of trust. *Currently under revision.*

†shared last authorship

Meyer, F., Engelmann, J.B., Fehr, E., Ruff, C.C. Impulsivity, sensation seeking and negative emotionality relate to distinct morphometric brain features**.** *In preparation.*

Contents

1 General Introduction

Decades of economic research have denied the essential role of emotion in decision-making and choice. That is, standard economic models assume that human decision-making is consequentialist in nature: Decision-making is postulated to be guided by the decision-maker's rational assessment of desirability and likelihood of alternative outcomes, i.e., by his strive to maximize utility (Rick & Loewenstein, 2008). Expected utility (EU) theory, a canonical normative model of decision-making under risk and uncertainty, is such a normative benchmark defining the axioms that are necessary and sufficient for the representation of an individual's preferences by means of an expected utility function (e.g., von Neumann & Morgenstern, 2007).

However, empirical evidence has demonstrated systematic violations of EU theory. Importantly, researchers regularly found EU-inconsistent risk aversion patterns such as the reflection effect (Kahneman & Tversky, 1979), which describes risk aversion for potential gains compared to risk-seeking for potential losses, and set out to explain these deviations. It has been argued that the observed behavioral anomalies can be ascribed to the unrealistic assumptions EU theory makes about emotions (Loewenstein & Lerner, 2003). Therefore, influential approaches such as the Appraisal-Tendency-Framework (ATF; Han, Lerner, & Keltner, 2007; Lerner & Keltner, 2000, 2001) aim to explain these deviations by more realistic assumptions about the influence of emotional factors on decision-making and judgment. The ATF distinguishes between the distinct effects of specific emotions on decision-making, assuming that different emotions elicit specific motivational and cognitive processes. It predicts domain-specific effects of emotion on decision-making due to emotion-specific appraisal tendencies. According to the ATF, anxiety is associated with the appraisal of an uncertain existential threat and situational control, resulting in an action tendency to reduce risk in anxious states. Anger, in contrast, is characterized by individual control and certainty, leading to more optimistic risk estimates and risk-seeking choices. Emotional states also influence the depth of processing. It has been shown, for instance, that high certainty emotions such as anger and happiness lead to heuristic processing, whereas low certainty emotions such as fear and sadness are characterized by analytic processing (Tiedens & Linton, 2001)

Together with neuroscientific findings on affective influences on decision-making (for review, see Phelps, Lempert, & Sokol-Hessner, 2014), models such as the ATF called for the integration of emotion into economic models of decision-making and the inclusion of emotion-related factors into utility calculation.

Emotion is not a unitary concept, but can broadly be defined as a discrete and time-restricted response to internal or external events, comprising features such as subjective experience, physiological responses, bodily expression and action tendencies (Phelps, 2009). Emotional factors have been increasingly acknowledged to influence humans' decision-making. We can distinguish these multifaceted, biologically mediated emotional responses, which are elicited by specific internal or external events, from affect, which is an umbrella term comprising emotion, mood and emotion-related traits. Mood, in contrast to emotion, is primarily marked by subjective feelings without necessary psychophysiological changes, a lack of a contextual "trigger" and longer durations. A common taxonomy of emotion employed in psychology and behavioral economics is the distinction between *expected* and *immediate* emotions: Expected emotions, on the one hand, are expected to be experienced as a future consequence of choices taken. Importantly, these predicted emotions are not experienced at the moment of choice, but materialize during the experience of the consequences of the selected option. Their inclusion into the calculation of expected utility explained aberrations from EU theory's axiomatic predictions and improved the predictive power of economic models of decision-making. For instance, changes in probability weighting could be predicted by consideration of "affective utility" (Walther, 2003). Immediate emotions, on the other hand, are experienced at the moment of choice and only compatible with classic economic theories when they are normatively *related* to the decision-making situation at hand (*integral* immediate emotions). *Incidental* immediate emotions are baseline affective states that can be situational or dispositional (e.g., personality or affective traits). They are normatively *unrelated* to the decision-making situation at hand. Hence, the influence of incidental emotion on choice and judgment represents a major violation of classic economic theories of human decision-making, as this would imply that decisions are influenced by factors unrelated to the utility of their consequences (Rick & Loewenstein, 2008).

The influence of incidental emotions on choice has thus been denied by traditional economic theory. Psychological research, however, has found that incidental emotions carry over to the subjective value assessment of available choice options and emphasized the relevance of incidental emotional factors for the predictive power of economic models

(Loewenstein & Lerner, 2003; Rick & Loewenstein, 2008). A multitude of studies showed that emotions elicited by stimuli unrelated to the decision-making situation at hand influence judgment and cognitive processing (e.g., Bodenhausen, Sheppard, & Kramer, 1994; Schwarz & Clore, 1983). People's judgments are often related to their current affective states, influencing, for example, consumer judgments (Yeung & Wyer, Jr., 2004), life satisfaction judgments (Schwarz & Clore, 1983) and economic decisions (Lerner, Small, & Loewenstein, 2004). Furthermore, individual differences in affective and personality traits are systematically associated with interindividual variation in decision-making. Studies have revealed, for instance, a positive relationship between trait anxiety and risk-avoidant decision-making; normally and pathologically anxious people typically tend to take less risk in choice situations (e.g., Giorgetta et al., 2012; Grecucci et al., 2013; Maner et al., 2007). Empirical findings have paved the way for the rising acknowledgment of incidental emotion in behavioral economics. However, the specific behavioral and neurobiological mechanisms leading to the modulatory effect of incidental emotions on decision-making are largely unknown (Lerner, Li, Valdesolo, & Kassam, 2015; Phelps et al., 2014).

Neuroeconomics, a discipline investigating the neural basis of decision-making by means of combining tools from neuroscience, economics and psychology, has recently been used to gain knowledge on decision-making processes in healthy persons (Platt & Huettel, 2008; Rangel, Camerer, & Montague, 2008) and on aberrant decision-making processes in psychiatric disorders (Kishida, King-Casas, & Montague, 2010). It offers quantitative and objective assessment of both behavioral and neural decision-making parameters and supports endeavors to identify biomarkers of normal and pathological cognition, thus helping to identify the neurobiological basis of psychiatric diseases. For instance, as affective and anxiety disorders are characterized by both deviations in decision-making and affect, neuroeconomic research has recently started to identify the relationship between cognitive and affective aberrations in those disorders (e.g., Huys, Pizzagalli, Bogdan, & Dayan, 2013; Paulus & Yu, 2012).

In line with this, further insights into the behavioral and neurobiological mechanisms that determine how specific emotions influence and interact with the decision-making process may be especially valuable for psychiatric research and diagnostics of maladaptive emotional responses that require a deeper knowledge on the influence of particularly incidental emotions (Kishida et al., 2010; Kishida & Montague, 2013).

Hence, this dissertation project set out to shed light on the behavioral and neurobiological underpinnings of emotion-related influences on decision-making in order to bridge the persistent knowledge gap on the interaction between affect and decision-making outlined above. The remainder of this dissertation is organized as follows:

First, I will summarize relevant theoretical approaches from psychology and economics to the interaction of emotion and decision-making followed by behavioral and neurobiological findings on the modulatory role of emotion in decision-making. Second, I will provide an overview of the motivation and rationale underlying the four studies I conducted (Chapter 1). Chapter 2 comprises detailed summaries of these studies. Based on the summaries, I discuss these studies, the deducible conclusions, and ideas for future studies in Chapter 3. The original papers discussed in this dissertation can be found in Appendix A, B, C and D.

1.1 Emotion-related influences on decision-making: Past research and theory

1.1.1 Integral and incidental emotional influences on decision-making

Both psychology and traditional economic theory acknowledge the influence of integral emotions on preferences. However, the disciplines disagree about the role of incidental emotions in decision-making. Whereas psychology has acknowledged the influence of incidental emotions on decision-making, these task-unrelated dispositional (e.g., affective traits) and situational (e.g., affective states) emotional factors and their influence on choice and judgment have been denied by traditional economic theory due to the axiomatic properties of original EU theory. Nevertheless, in recent years, the accumulating empirical evidence for the modulatory influence of incidental emotions on decision-making could not be ignored by economic theory either. The following sections provide an overview of theoretical and empirical aspects of integral and incidental emotions with respect to their influence on decision-making, with a focus on yet unanswered questions.

As pointed out by several researchers (e.g., Tversky & Kahneman, 1974; Walther, 2003; Xu et al., 2013), emotions may influence decision-making via choice heuristics that are informed by affective responses to the decision-making context, especially when confronted with potentially aversive outcomes. It was found that the anticipation of

potential outcomes of a risky decision evokes emotional reactions and influences valuation and preferences (Caplin & Leahy, 2001; Loewenstein, 1987). These "anticipatory" emotions, which are inherent to or elicited by the choice situation, often propel behavior in directions different from those predicted by EU theory. There are two common behavioral deviations from EU theory, which have stimulated abundant research in economics, psychology, and neuroeconomics and are often attributed to the influence of *integral* emotions on decision-making.

The first of the behavioral phenomena describing a stark deviation from rational choice theory is the *framing effect*. It is a strong violation of the description invariance axiom of EU theory implying that equivalent descriptions of a decision problem lead to systematically different choices. That is, discrete choices between a risky and a riskless option of equal expected value depend on the description of the decision problem in positive versus negative terms ("gain frame" versus "loss frame", respectively). It has been repeatedly observed that positively framed descriptions of such choice sets are associated with risk-aversion, whereas negatively framed choice sets elicit risk-seeking (e.g., Kühberger & Tanner, 2010; Tversky & Kahneman, 1986). The framing effect has first been explained by prospect theory (PT; Kahneman & Tversky, 1979), which is a descriptive model of decision-making under risk. PT is characterized by several core components: It comprises an inverse-S-shaped probability weighting function, w(p), capturing overweighting of low probabilities and underweighting of moderate to high probabilities. It also comprises a value function, v(x), over losses and gains relative to a reference point. The value function is convex for losses and concave for gains, contributing to risk-aversion for gains and risk-seeking for losses, respectively. It is assumed that the framing manipulation determines whether outcomes are evaluated in terms of gains or losses relative to a reference point (often the status quo), thus eliciting distinct risk-averse or risk-seeking behavioral patterns.

The investigations of the basis of the framing effect resulted in its conceptualization as an emotional bias in decision-making, which is attributed to the influence of integral emotions on risky choice (e.g., Cassotti et al., 2012; De Martino, Kumaran, Seymour, & Dolan, 2006; Druckman & McDermott, 2008; Xu et al., 2013). The emotional account of the framing effect is supported by neuroscientific evidence. Functional neuroimaging studies have investigated the neural correlates of different framing effect scenarios. Most prominently, De Martino et al. (2006) investigated the susceptibility to framing in risky economic choice. These researchers mainly compared neural activity for frame-congruent

choices with frame-incongruent choices. That is, they contrasted neural activity during choices made according to the behavioral tendency imposed by the frame (i.e., choosing the safe option in the gain frame versus choosing the risky option in the loss frame) with neural activity for choices taken against the behavioral tendency imposed by the frame (i.e., choosing the risky option in the gain frame versus choosing the safe option in the loss frame). While increased amygdala activation was observed for frame-congruent decisions, frame-incongruent decisions were correlated with increased activation in dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (ACC). This activation pattern, especially the amygdala activation for frame-congruent decision-making, was interpreted as evidence for an emotion-driven framing effect, which can be overcome by the implementation of cognitive control over behavioral tendencies imposed by the frames via DLPFC and ACC (however, see Talmi, Hurlemann, Patin, & Dolan, 2010 for conflicting results). Strong empirical evidence demonstrates the substantial role of both ACC and DLPFC in the implementation and maintenance of top-down cognitive control over behavior and decision-making (Coutlee & Huettel, 2012; MacDonald, Cohen, Stenger, & Carter, 2000; Miller & Cohen, 2001). Both right and left DLPFC have been observed to be involved in computations relevant to those functions (Fecteau et al., 2007; Figner et al., 2010; Hare, Camerer, & Rangel, 2009; Knoch et al., 2006). However, direct causal evidence for the role of prefrontal brain structures in resistance to framing is missing since the conducted neuroimaging studies (De Martino et al., 2006; Gonzalez, Dana, Koshino, & Just, 2005; Xu et al., 2013) offer solely correlational evidence of brain structures with functional relevance for framing effects. Hence, an endeavor of this dissertation project was the investigation of the neural mechanisms causally underlying the impact of and resistance to integral emotions on decision-making.

The second important deviation from the predictions of EU theory is *loss aversion*. It describes the consistently observed increased sensitivity to losses compared to gains and people's tendency to place greater weight on potential losses than equal-sized gains. Loss aversion is another core characteristic of PT and captured by the value function's greater steepness for losses than gains. Although no canonical definition of loss aversion exists, the loss aversion coefficient, λ, is usually calculated as the ratio between the slope of the value function for losses and the slope of the value function for gains, that is, the mean or median of $v'(-x)/v'(x)$ (Fox & Poldrack, 2009). It is suggested that more hedonic impact is lent to losses than gains leading to multiplicative overweighting of losses relative to gains.

Camerer (2005) has even described loss aversion as a "fearful overreaction to transitory loss" emphasizing the emotional component hypothesized to underlie the aversion to losses.

Electrophysiological and neuroimaging studies also set out to illuminate the neurobiological mechanisms underlying loss aversion. Sokol-Hessner et al. (2009) observed increased skin conductance responses (SCRs) to losses relative to gains: This finding was interpreted as evidence for the increased emotional arousal associated with losses relative to gains as increases in SCR are an indicator of increased emotional arousal. SCR measurement has been successfully used in previous research to monitor and validate changes in emotional arousal (e.g., Grupe & Nitschke, 2011). Furthermore, research demonstrated that the degree of loss aversion can be manipulated by means of emotion regulation techniques (Sokol-Hessner et al., 2009; Sokol-Hessner, Camerer, & Phelps, 2013) and that the observed reduction of loss aversion due to emotion regulation correlates with decreased amygdala activation during the choice situation (Sokol-Hessner et al., 2013). These findings supported the conceptualization of loss aversion as an emotion-driven process and were interpreted as empirical evidence for the impact of integral emotions (i.e., negative emotions inherent to losses) on decision-making. In a similar vein, De Martino, Camerer, and Adolphs (2010) demonstrated decreased degrees of loss aversion in patients suffering from Urbach-Wiethe disease (UWD), which is characterized by amygdala calcification, and thus draw the conclusion that loss aversion is driven by an emotional mechanism. However, these results should be taken with care as we cannot rule out the existence of other neural dysfunctions associated with UWD in those patients, and the possibility of confounding anatomical deviations limits the explanatory power of lesion studies in general. Contrary to the above described findings, another study on the neural correlates of loss aversion (Tom, Fox, Trepel, & Poldrack, 2007) did not find loss- and loss aversion-related activations in amygdala and insula, which are brain regions typically associated with emotional processes and the processing of negative emotional stimuli (LeDoux, 2003; Paulus & Stein, 2006). Tom et al. (2007) rather pointed to a one-dimensional valuation system processing both gains and losses instead of separate neural coding systems for losses versus gains. They concluded that loss aversion is not very likely to be an emotional overreaction related to fear. Importantly, however, this conclusion has to be taken with care as the authors drew it upon a null result and based on reverse inference.

To sum up, loss aversion reflects valuation-related processes and might be used to identify changes in valuation and subjective value assessment invoked by changes in incidental emotional states. Although accumulating empirical evidence supports the notion

of loss aversion as an emotional phenomenon (De Martino et al., 2010; Sokol-Hessner et al., 2009, 2013; Sokol-Hessner, Hartley, Hamilton, & Phelps, 2014), a final conclusion about the behavioral and neurobiological processes underlying loss aversion has not yet been drawn.

Taken together, both loss aversion and the framing effect may be characterized as being based on the effect of integral emotions on decision-making, as these behavioral phenomena represent individual emotional responses that are elicited by the characteristics of a choice at hand. However, we still face deficient knowledge on the causal neural mechanisms leading to the effect of integral emotions on choice and judgment. I therefore conducted a study to investigate the neural mechanisms causally underlying framing effects on decision-making.

On the other hand, task-unrelated emotional influences on decision-making can be described as emotional background to the decision at hand, without any normative relation between this baseline affective state and the decision-making task. These incidental emotional responses are typically elicited by situational sources (e.g., sunny versus rainy weather) or due to dispositional sources (e.g., dispositional anxiety). In recent years, a multitude of studies showed the influence of incidental emotions on decision-making (e.g., Andrade & Ariely, 2009; Harlé & Sanfey, 2007; Lerner & Keltner, 2001; Lerner et al., 2004). Different theoretical approaches have been developed to explain the mechanisms underlying incidental (but also integral) emotional influences on judgment and choice. Early valence-based approaches categorizing affect into positive and negative dimensions stated that emotions of the same valence should have similar effects on decision-making and risk perception (Johnson & Tversky, 1983). Although parsimonious, the valence-based approach was not able to explain the observed opposing effects of emotions of similar valence on decision-making (for review, see Loewenstein & Lerner, 2003). Based on these findings, models taking into account specific effects of emotion on judgment and choice such as the ATF (Lerner & Keltner, 2000) have been developed. As described above, it is assumed that specific emotions are linked to cognitive appraisals and motivational goals, which in turn are associated with specific action tendencies. By means of these action tendencies and associated goal-directed processes, emotions exert influence on decision-making. The validity of ATF predictions has been verified in a series of experiments (e.g., Han et al., 2007). According to the ATF, the effects of emotion on decision-making are mediated by emotion-specific appraisals of certainty and control rising from specific emotions. Despite existing psychological theories and rich empirical evidence, economic

research still fails to include the influence of incidental emotions in its theories of human decision-making (Loewenstein & Lerner, 2003; Rick & Loewenstein, 2008).

Apart from theoretical considerations, we can clearly observe a newly developed interest in the mechanisms underlying the effects of emotions on judgment and choice – with a focus on negative emotions such as anxiety. This is evident from the large number of studies conducted on this issue (e.g., Clark et al., 2012; Giorgetta et al., 2012, 2012; Robinson, Charney, Overstreet, Vytal, & Grillon, 2012; Robinson, Letkiewicz, Overstreet, Ernst, & Grillon, 2011; Robinson, Overstreet, Charney, Vytal, & Grillon, 2013). Acute or state anxiety represents an adaptive emotion and carries substantial evolutionary value as a functional response to imminent danger, but pathological degrees of anxiety imply detrimental consequences for individual functioning (Rosen & Schulkin, 1998). Furthermore, anxiety forms a major personality trait, referring to an individual's relatively stable disposition to feel anxious, i.e., his general level of anxiety. Pathological anxiety as found in psychiatric diseases such as anxiety and affective disorders is associated with aberrant decision-making (Monterosso, Piray, & Luo, 2012; Paulus & Yu, 2012; Treadway & Zald, 2011). Both dispositional anxiety and state anxiety generally correlate with increased risk avoidance (for review, see Blanchette & Richards, 2010). Facets of dispositional anxiety – trait anxiety, worry, and social anxiety – correlate with increased risk-avoidant decision-making in the Balloon Analogue Risk Task (BART) in healthy subjects (Maner et al., 2007). This pattern of increased risk-avoidant behavior was also found in pathological trait anxiety (Giorgetta et al., 2012; Maner et al., 2007) (see section 2.3). In line with the findings on the effects of dispositional anxiety on decision-making, investigations employing the experimental induction of anxious states also report generally increased risk-avoidant behavior in anxious or stressed states (Clark et al., 2012; Maner et al., 2007).

Despite the recent awareness of the substantial interaction between anxiety and decision-making along with the detrimental consequences of pathological anxiety (Grupe & Nitschke, 2013; Hartley & Phelps, 2012), we still face deficient knowledge on the neurobiological substrates of potential interactions between incidental anxiety and decision-making. The trajectory from normal to pathological anxiety is thought of as a continuum (Rosen & Schulkin, 1998) and insights into neural processes underlying the progression from adaptive to maladaptive anxiety may bear implications in regard to treatment and prevention of psychiatric disorders. To the best of our knowledge, no combined behavioral-functional studies on the interaction between emotion and decision-making have been

published yet. This seems especially surprising considering that both normal and pathological anxious states cause aberrant decision-making in complex social and non-social situations (e.g., Clark et al., 2012; Grecucci et al., 2013). Furthermore, taking into account the neural correlates of both economic decision-making and anxiety, substantial interactions between the two domains seem very likely as a broad overlap in the brain regions involved in both emotion and economic valuation exists (for review, see Hartley & Phelps, 2012). Brain structures such as the amygdala, insula, ventral striatum (VS) and ventromedial prefrontal cortex (VMPFC) have been shown to be involved in both emotion processing and value-based decision-making (Chib, Rangel, Shimojo, & O'Doherty, 2009; De Martino et al., 2010; Denny, Ochsner, Weber, & Wager, 2013; Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005; Levy & Glimcher, 2012; Rangel et al., 2008). Despite this obvious overlap between the neurocircuitries of anxiety and decision-making, substantial knowledge about the interaction between incidental emotion- and economic valuation-related neural processes is lacking.

In order to elucidate the impact of adaptive and maladaptive anxiety on decision-making with regard to their neural and behavioral underpinnings, I set out to add to the small number of studies conducted on this issue in the behavioral (e.g., Berghorst, Bogdan, Frank, & Pizzagalli, 2013; Clark et al., 2012) or neurofunctional domain (e.g., Gold, Morey, & McCarthy, 2014). I thus investigated the influence of experimentally induced anxiety on decision-making with two functional magnetic resonance imaging (fMRI) studies.

1.1.2 Social and non-social frameworks of decision-making

Emotions are said to be inherently social, that is, they are often elicited by and expressed towards other people. Humans are also able to regulate their emotions in order to adapt to social norms and influence other persons (Fischer & Manstead, 2009; Keltner & Lerner, 2010; Kleef, 2009). Thus, it seems important to also consider the interaction between emotional and social factors in decision-making. We can distinguish between situations where choices take place in a "social vacuum", that is, individual choices in which the decision-maker must only consider his own values and preferences to select an option, and situations in which the individual's outcome depends on concomitant choices of others. A behavioral economic paradigm comprising this social component is, for instance, the trust game (Berg, Dickhaut, & McCabe, 1995), while the classic mixed gamble task (e.g., Tom et al., 2007) is an experimental paradigm of individual (i.e., non-social) decision-making.

Importantly, recent behavioral and neurofunctional studies point to the substantial influence of incidental emotions on social decision-making situations. For example, experimentally induced incidental disgust has been shown to modify selling and buying prices (Lerner et al., 2004), behavior in the ultimatum game (Bonini et al., 2011; Harlé & Sanfey, 2010) and moral decision-making (Inbar, Pizarro, Knobe, & Bloom, 2009; Schnall, Haidt, Clore, & Jordan, 2008). Furthermore, it has been demonstrated that incidental sadness reduces selling prices, but increases buying prices, leading to a "reversed" endowment effect (Lerner et al., 2004). Sadness has also been associated with increased dictator generosity in the dictator game (Tan & Forgas, 2010). Harlé & Sanfey (2007) reported increased rejection rates of unfair offers in sad as opposed to amused and neutral mood states. Although multiple studies compared social to non-social decision-making (e.g., Phan, Sripada, Angstadt, & McCabe, 2010; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003), research on the neural correlates of emotional influences on social decision-making is rare (however, see Harlé, Chang, van't Wout, & Sanfey, 2012). One may speculate about interactions between emotion- and social cognition-related neural computations for social decision-making situations, which are not present in non-social decision-making tasks (for a framework of social versus non-social decision-making, see Ruff & Fehr, 2014). Thus, highly specific modulatory effects of emotion on social versus non-social decision-making may be expected. To probe this conjecture, I conducted two fMRI studies, which investigated the influence of incidental anticipatory anxiety on decision-making in a non-social and a social context decision-making task.

1.1.3 Influence of personality and emotion-related traits on decision-making

In recent years, the modifying influence of personality and affective traits on various forms of decision-making has been increasingly acknowledged (Davis, Patte, Tweed, & Curtis, 2007; Giorgetta et al., 2012; Lauriola & Levin, 2001; Lauriola, Panno, Levin, & Lejuez, 2013). Thus, a further study of this dissertation project aimed to complement research on the morphometric brain properties that uniquely predispose an individual towards a particular decision-making style.

The fact that no canonical definition of personality exists hampers research on associated topics, but most biopsychological theories state that personality reflects stable emotional and motivational attributes of a person (DeYoung & Gray, 2009; Yarkoni, 2015).

Given that the proximal source of human behavior is the brain, empirical evidence demonstrated that variation in brain anatomy and function correlates with interindividual variation in behavior (Kanai & Rees, 2011). Decades of research revealed the heritability and genetic basis of personality traits and identified the stable structure of personality to be based on biological factors (Bouchard Jr & Loehlin, 2001; Van Gestel & Van Broeckhoven, 2003). fMRI studies reported that individual personality differences are reflected by differences in functional brain activity (e.g., Canli et al., 2001; Kumari, Ffytche, Williams, & Gray, 2004). Thus, researchers became interested in complementing the knowledge about the biological basis of personality and set out to expand it to the neuroanatomical dimension. The relevance of personality was additionally emphasized by its observed modulatory effects on decision-making, which has been repeatedly demonstrated (Lauriola & Levin, 2001; Lauriola et al., 2013).

The study I conducted focuses on the neurobiological basis of stable personality and affective traits that are commonly assumed relevant to individually varying decision-making styles. The aim was to establish a comprehensive assessment of those traits relevant for decision-making in combination with brain morphometric data. Personality and affective traits, which have consistently been shown to modulate decision-making, were targeted in this investigation. In the following paragraphs, these traits will be described.

Sensation seeking has been defined as the tendency to seek intense experiences. In order to experience varied, novel and stimulating sensations, sensation seekers are willing to take "physical, social, legal and financial risks" (Zuckerman, 1994). Sensation seeking is mostly assessed by means of self-report measures, which acknowledge sensation seeking as a heterogeneous construct. Zuckerman's Sensation Seeking Scale (SSS-V; Zuckerman, 1996) is an operational measure of sensation seeking tendencies assessing the personality construct via four subscales (thrill and adventure seeking, experience seeking, boredom susceptibility, disinhibition) designed to reflect the multiple dimensions of sensation seeking.

Impulsivity, on the other hand, is a key concept in psychopathology and conceptualized as deficient planning ability and lack of reflectiveness, carelessness, rapid action taking and decision-making. Impulsivity is not a unitary trait, but a complex multidimensional construct. Different measures have been developed to assess impulsivity. Stop-signal reaction time tasks (Logan, 1994) and Go/No-go tasks (Ruchsow et al., 2008) are commonly used in order to assess the inhibition of prepotent responses to assess facets of motor impulsivity. Delay discounting tasks (Rachlin & Green, 1972) have been used to

experimentally measure impulsivity in financial decision-making situations. Furthermore, self-report measures of impulsivity have been applied in order to quantitatively assess a person's subjective perception of his or her impulsive characteristics. The Barratt Impulsiveness Scale (BIS-11; Patton, Stanford, & Barratt, 1995) is frequently used to assess distinct aspects of individual differences in impulsive traits (motor, attention and non-planning dimension). Importantly, the different impulsivity measures assess distinct aspects of impulsivity and do not necessarily substantially correlate with each other (Bevilacqua & Goldman, 2013).

Both sensation seeking and impulsivity lead to interindividual differences in both real-world decision-making and experimental settings. Two studies (Sweitzer, Allen, & Kaut, 2008; Zermatten, Van der Linden, d' Acremont, Jermann, & Bechara, 2005) found that increased degrees of impulsivity led to increased disadvantageous decision-making in the Iowa gambling task. Furthermore, a meta-analysis (Lauriola et al., 2013) on the influence of impulsivity versus sensation seeking on behavior in the Balloon Analogue Risk Task (BART) identified small to moderate effect sizes for the association between sensation seeking and risky decision-making, whereas only small effect sizes for the relation between impulsivity and risk-taking were found. Despite the conceptual relatedness between sensation seeking and impulsivity, prior research demonstrates the need to distinguish between these two heterogeneous concepts, as different associations of impulsivity and sensation seeking, for example, with addiction, were reported and their contribution to decision-making in health and psychiatric diseases seems to differ (Ersche, Turton, Pradhan, Bullmore, & Robbins, 2010). Furthermore, factor analysis suggested impulsivity and sensation seeking to be different concepts (Magid, Maclean, & Colder, 2007).

Lastly, negative emotionality describes the tendency to experience negative emotional states. It corresponds to a dominant neuroticism factor within the Big Five personality inventory (Costa & McCrae, 1992) and has been shown to be related to affective and anxiety disorders (Boschloo et al., 2013; Tackett, Waldman, Van Hulle, & Lahey, 2011). It was demonstrated that pathological negative affectivity (i.e., increased probability of experiencing negative moods and emotions as well as emotional distress as seen in various psychiatric disorders) is associated with changes in decision-making (Giorgetta et al., 2012; Grecucci et al., 2013). As described above, several behavioral studies have shown that increased trait anxiety, a component of negative emotionality, is associated with increased risk avoidance (Giorgetta et al., 2012; Maner et al., 2007). Excessive anxiety, in

turn, is a core symptom of and risk factor for different psychiatric disorders (Grupe & Nitschke, 2011).

Taking into account the heritability of personality and genetic factors moderating interindividual variation in impulsivity, sensation seeking, and negative emotionality (Bevilacqua & Goldman, 2013; Domschke & Deckert, 2012; Stein, Jang, & Livesley, 1999; Stoel, De Geus, & Boomsma, 2006), it seems very likely that there are structural and functional brain correlates of those personality and affective traits as well.

In fact, several studies have investigated the structural as well as functional correlates of impulsivity, yielding an inconsistent picture of the neural representation of impulsivity. Schilling et al. (2013) measured impulsivity by means of the Temperament and Character Inventory (TCI-R; Cloninger, 1999) and found a negative relationship between impulsivity and left orbitofrontal (OFC) grey matter (GM) volume in a healthy adolescent sample. Matsuo et al. (2009) also reported reduced left and right OFC GM volume as a function of increasing impulsivity (total BIS-11 score). More specifically, these authors observed a negative relationship between motor impulsivity and left OFC as well as a smaller right OFC GM volume in case of higher non-planning impulsivity. Furthermore, a negative relationship between anterior cingulate volume and impulsivity (total BIS score) was observed. In contrast, another study (A. K. W. Lee, Jerram, Fulwiler, & Gansler, 2011) reported negative associations between attention and motor impulsivity with left superior temporal gyrus and a positive relation between non-planning impulsivity and left OFC GM volume and lateral frontopolar cortex. These inconsistent results are most likely due to methodological issues, such as personality measures employed, analysis methods used for behavioral and brain morphometric data (e.g., simple versus multiple regression) and the number of subjects.

In comparison, studies on the neuroanatomical correlates of sensation seeking are rare. Martin et al. (2007) reported a positive relationship between right hippocampal volume and experience seeking, a facet of sensation seeking. A study on the functional correlates of sensation seeking (Joseph, Liu, Jiang, Lynam, & Kelly, 2009) found that individuals with strong sensation seeking tendencies showed stronger fMRI responses to high-arousal stimuli in insula and posterior medial orbitofrontal cortex, whereas low sensation seeking persons were characterized by greater activation to high-arousal stimuli in anterior cingulate cortex and anterior medial orbitofrontal regions. These results have been interpreted as an overactive approach system in high sensation seekers versus a dominating inhibition system in low sensation seekers.

Regarding negative emotionality, different components of this personality construct were investigated. For instance, researchers sought to identify the neuroanatomical correlates of state and trait anxiety (Spampinato, Wood, De Simone, & Grafman, 2009) using the state-trait-anxiety inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). While they found a positive correlation between anxiety and ventrolateral prefrontal cortex (inferior frontal gyrus), inverse correlations between anxiety and several brain regions such as bilateral dorsolateral prefrontal cortex, rostral anterior cingulate cortex and post/retrosplenial cingulate cortex have been observed. Yamasue et al. (2008) employed the TCI-R in order to assess harm-avoidance as an anxiety-related measure and found gender-unspecific reduced right hippocampal volume as a function of increasing anxiety. Furthermore, they reported a negative correlation between anxiety and left anterior prefrontal cortex GM volume in female but not male participants. A further study (Omura, Constable, & Canli, 2005) found a negative correlation between anxiety (assessed via NEO-PI-R neuroticism scale) and right amygdala GM volume, while DeYoung et al. (2010) observed negative relationships between the same personality measure and dorsomedial prefrontal cortex, medial temporal lobe and precentral gyrus. Positive associations were observed for mid-cingulate cortex, caudate, and regions in middle temporal gyrus and cerebellum. Notably, Hu et al. (2011) could not find significant correlation between neuroticism and local GM volume. Although most studies identified associations between interindividual variation in anxiety and anxiety-related measures and brain structure, the variation in results across studies is remarkable. Different personality measures, analysis methods, and numbers of subjects are again possible explanatory factors for the inconsistent results yielded in these studies.

In conclusion, improved assessment of personality and affective traits, its combination with brain morphometric data, and an appropriate number of subjects may be necessary in order to achieve valid results on the neuroanatomical basis of personality and its modulatory role in interindividually varying decision-making styles.

1.2 Purpose of this dissertation

Although recent research has identified neurobiological correlates of decision-making on the one hand and has investigated behavioral and neurofunctional properties of emotions and personality on the other hand, we still lack detailed knowledge about the neural

patterns, which underlie the interaction between decision-making and emotion-related factors. Furthermore, the context of the decision at hand, in the sense of a social or non-social context the decision is made in, and potentially systematic interactions with emotional factors deserve further attention and clarification. Therefore, I set out to a more nuanced investigation of these relationships. The following four investigations on the impact of affective factors on decision-making were conducted:

(1) I performed a transcranial magnetic stimulation (TMS) study on the causal mechanisms underlying inhibition of risky choice framing effects (i.e., suppression of affective heuristics, which are supposed to be based on the effect of integral emotion on decision-making (section 2.1 and Appendix A).

(2) The behavioral and neurofunctional effects of incidental emotions on risky economic decision-making in a non-social context were investigated in an fMRI study where incidental anticipatory anxiety was induced via the threat-of-shock technique. The effects of incidental anticipatory anxiety on subjective value assessment were measured with behavioral, electrophysiological and neuroimaging tools (section 2.2 and Appendix B).

(3) The behavioral and neurofunctional effects of incidental anticipatory anxiety on social decision-making were investigated in a second fMRI study with the main goal to investigate changes in trusting behavior due to anticipatory anxiety using a within-subject design (section 2.3 and Appendix C).

(4) Lastly, the neuroanatomical basis of personality and affective traits relevant to interindividual variation in decision-making styles was investigated employing voxel-based morphometry (VBM). This study was also meant to address the colinearity of personality constructs on the behavioral and neuroanatomical level (section 2.4 and Appendix D).

2 Contents of this dissertation

2.1 Study A: Framing effects on risky decision-making: Investigating a causal role for the dorsolateral prefrontal cortex with TMS

I conducted a transcranial magnetic stimulation (TMS) study together with Daniela M. Pfabigan and Christian C. Ruff in order to examine the neural mechanisms causally underlying the effect of integral emotions and associated affective heuristics on decision-making. Motivated by an fMRI study of De Martino et al. (2006), which provided evidence for the inhibitory role of DLPFC in risky choice framing effects (Tversky & Kahneman, 1986), I examined whether the DLPFC is causally linked to the inhibition of the framing effect. The framing effect is hypothesized to reflect emotional biases in decision-making and describes people's tendency to react to the same choice problem in different ways depending on whether it is presented in terms of losses (*loss frame*) or gains (*gain frame*). The DLPFC is a brain region well known for its involvement in cognitive control processes and emotion regulation (e.g., MacDonald et al., 2000; Phillips, Ladouceur, & Drevets, 2008). Furthermore, bilateral DLPFC was identified as neural correlate of frame-inconsistent decisions in recent neuroimaging studies. That is, De Martino et al. (2006) found significantly increased BOLD signal for trials in which subjects were able to overcome the general behavioral tendencies imposed by the frames. Similar results were obtained by Gonzalez et al. (2005) reporting increased DLPFC activation for risky relative to sure choices in the positive frame employing the classic Asian disease Problem (Tversky & Kahneman, 1986). Overcoming a behavioral heuristic seems to require additional cognitive resources, and the correlative fMRI results suggest that overcoming frames necessitates recruitment of the DLPFC. Based on the hypothesis that the DLPFC is causally involved in inhibition of the framing effect, we hypothesized to find increased susceptibility to framing due to the temporary disruption of DLPFC activity.

I employed TMS in order to achieve a temporary "virtual lesion" of DLPFC. Immediately following the applied continuous theta burst stimulation (cTBS, a TMS

protocol leading to the temporary dysfunction of the target brain region) of either unilateral DLPFC (left or right [experimental conditions]) or vertex (control condition), the subjects had to perform a risky choice framing task adapted from De Martino et al. (2006). The decision-making task consisted of repeated choices between a lottery and an expected value-equivalent sure option, which was critically framed as a gain or loss. After subjects were endowed with varying amounts of money, they had to decide whether to play a lottery in order to win or lose the whole endowment amount (with varying probabilities of winning [p(win)] and losing [p(lose) = 1- p(win)]) or to take a fraction of the endowment for sure, which was either expressed in terms of gains or losses (gain frame or loss frame, respectively).

Against expectations, the current results do not support the hypothesis of a causal role of DLPFC in inhibition of the framing effect. Although neuroimaging studies suggested a substantial role of DLPFC in framing, we could not confirm the functional relevance of the DLPFC for inhibition of the framing effect due to unilateral disruption of either left or right DLPFC in a financial risky choice framing task. No effect of left or right DLPFC activity disruption on the susceptibility to framing was observed. That is, preferences between stimulation groups did not significantly differ from the control group in the TMS session. Detailed statistical analysis of subjects' choices (employing summary statistics, linear and logistic regression analyses) disclosed a stable framing effect across all three groups without significant differences between groups. That is, although we find significantly more gambling choices in the loss frame relative to gain frame in all groups, disruption of DLPFC function did not lead to changes in susceptibility to framing, i.e., we did not observe significant changes in behavior as a function of TMS group. Examining interactions between TMS and different probability of winning levels and varying endowment amounts used in the gambles as well as controlling for various personality traits, we did also not find differential effects of unilateral DLPFC TMS on framing susceptibility. cTBS did not influence cognitive processing speed either, reflected in no reaction time differences between DLPFC and vertex groups. While increased reaction times were observed in the loss relative to gain frame, no significant reaction time differences between groups were observed.

Several conclusions and motivations for future studies can be drawn from this TMS study. On the one hand, it is possible that the causal role of DLPFC in inhibition of framing susceptibility had been overestimated due to the results of previous neuroimaging studies. On the other hand, modifications of the TMS approach employed in the current experiment

may be necessary in order to be able to draw conclusions on the functional relevance of DLPFC for risky choice framing. Based on our results in combination with previous neuroimaging studies on the framing effect, future studies should investigate the effects of bilateral DLPFC inhibition as well as dual-site TMS of DLPFC and ACC in order to identify brain regions causally involved in the generation and inhibition of risky choice framing effects. Specifically, based on the involvement of the DLPFC in frame-incongruent choices found in previous studies (De Martino et al., 2006; Gonzalez et al., 2005), we hypothesized about a causal role of DLPFC in the inhibition of behavioral tendencies imposed by frames. Taking into account our results, it may instead be speculated that DLPFC implements computations that are correlated with neural processes relevant for framing resistance, but these computations do not contribute to framing resistance itself. DLPFC activity might be a byproduct of frame-resistant ACC activity without any functional relevance for inhibition of the framing effect. In this case, we could define the relation between DLPFC and susceptibility to framing as correlational. Functional and effective connectivity analyses in combination with concurrent TMS-fMRI studies may be suited to test between the correlational versus causal role of DLPFC in framing effects. The suggested experimental and methodological modifications would be suited to identify framing-relevant neural connectivity patterns and their directions as well as to offer causal evidence for the neurobiological substrate of framing effects. Although DLPFC might be involved in inhibition of susceptibility to framing, it is possible that bilateral disruption of DLPFC via TMS or tDCS electrode montage is needed in order to disrupt the control function over framing of the DLPFC as resistance against the frame correlated with bilateral DLPFC activity in the study by De Martino et al. (2006). Concurrent TMS-fMRI might be suitable to investigate the effect of (bilateral) disruption of DLPFC function on framing-related neural activation and associated behavioral correlates in more detail. As stated above, we may hypothesize that framing susceptibility is regulated by synchronized activation of a network of brain regions involved in cognitive control processes. De Martino and colleagues' study (2006) suggests the ACC as candidate region of this network in addition to the DLPFC. Dual-site TMS (O'Shea, Taylor, & Rushworth, 2008) may be required for simultaneous interruption of DLPFC and ACC in order to change subjects' susceptibility to framing; partial disruption of the network might engage compensatory processes, resulting in stable behavior. Although the anatomical properties of the ACC require the use of a double-cone coil, which reduces the precision of the stimulation of the

target area, previous studies have suggested that effective TMS of ACC may be made possible by means of this coil type (Hayward et al., 2007).

Further technical issues should also find attention when considering our experimental results. We cannot rule out that our procedure was compromised by distance-dependent TMS effects and that cTBS effects were weakened by the location of the scalp-distant DLPFC target, as rapid decline in magnetic field strength with distance of the targeted cortical sites from scalp surface weakens actual TMS effects (Stokes et al., 2005). Although coil position was administered according to current technical guidelines (Rossi, Hallett, Rossini, Pascual-Leone, & Safety of TMS Consensus Group, 2009) and precisely controlled via our neuronavigation device, it is possible that changes in coil position relative to the gyrus may be necessary in order to optimize the magnitude of cortical stimulation. Coil position and head-coil angle may have to be changed to yield optimal cTBS application. Future studies may test such technical modifications to investigate whether DLPFC functioning can be interrupted more profoundly and whether this affects resistance to framing effects.

More generally, we have to point to the fundamental issue concerning the interpretation of null results. Although we did not find evidence for the causal role of DLPFC in inhibition of risky choice framing effects, we are not able to strongly conclude on the functional relevance of this brain region for framing effects. That is, our experimental design might not have been sensitive enough to detect the undoubtedly small effects of our experimental manipulation found for the difference in framing effects between control and experimental groups (difference in framing effects between vertex and left or right TMS group: $d = 0.004$ and $d = 0.047$, respectively). However, in the case of such weak effects as found in this study, it seems reasonable to conclude that previous studies may have overestimated the role of the DLPFC in overcoming framing susceptibility, and that right DLPFC is presumably not the main neural locus at which framing-related neural processes are causally implemented.

2.2 Study B: Anticipatory anxiety disrupts neural valuation during risky choice

This study, a conjoint research endeavor with Jan B. Engelmann, Ernst Fehr, and Christian C. Ruff, was conducted in order to examine the effects of incidental anticipatory

anxiety on risky decision-making and the associated neurofunctional correlates. The experiment was particularly designed to investigate the gap between psychological and economic theories on the interaction between incidental anticipatory anxiety and risky decision-making. We gathered behavioral, electrophysiological, and neural information in order to examine the influence of experimentally induced incidental anxiety on economic valuation under risk.

A main concern of this study was the employment of an improved emotion induction technique as the validity of emotion induction techniques used in previous studies is questionable with respect to the intensity, quality, and temporal properties of the induced emotion. That is, emotion induction techniques such as the cold pressor task (Porcelli & Delgado, 2009) or watching emotional movie clips (Lerner et al., 2004) require the affective manipulation to take place prior to the decision-making task instead of an "online" affective manipulation during the decision-making situation. The majority of previous investigations on the effects of incidental negative affective states such as stress on risk-taking employed the cold pressor task (e.g., Porcelli & Delgado, 2009; Porcelli, Lewis, & Delgado, 2012), which, for technical reasons, implies a between-subjects design. We therefore chose the induction of well-defined anticipatory anxiety via the translational threat-of-shock (TOS) technique (Robinson, Vytal, Cornwell, & Grillon, 2013; Schmitz & Grillon, 2012). Using the TOS paradigm, which creates a high incidental anxiety ("threat") condition and a low (or no) incidental anxiety ("safe") condition for each subject, we were able to investigate within-subject effects, i.e., individual changes in risk-taking due to the affective manipulation. By using a TOS paradigm we were able to investigate the effects of anticipatory anxiety on a trial-by-trial basis and elicit repeated, but transient emotional responses to threat, whereas, for instance, a cold pressor task relies on a single stressful event delivered to the subjects before the experimental task. Furthermore, the additional recording of skin conductance responses (SCRs) was chosen as a monitoring and validation tool of the affective manipulation, as the analysis of SCRs provides an index of the efficacy of anticipatory anxiety induction by the TOS manipulation. Therefore, we created a hybrid fMRI design, that is, the subjects' affective states were manipulated on a block-wise basis by means of the TOS manipulation, while they completed a risky choice task, which was based on an event-related trial design, in an fMRI environment.

In order to examine the effects of incidental anticipatory anxiety on individuals' choices and associated subjective value processing, we employed an economic paradigm that is well-established in behavioral economics. We adapted a mixed gamble task, which

has been used in previous functional neuroimaging studies (Minati, Grisoli, Franceschetti, et al., 2012; Minati, Grisoli, Seth, & Critchley, 2012; Tom et al., 2007) to assess changes in the behavioral and neural correlates of economic valuation due to the affective manipulation. Subjects repeatedly chose between an equiprobable lottery option, which comprised a gain and a loss outcome, and a sure option. The outcome of the sure option was fixed to a value of zero for all trials. Using this task, we were able to observe potential changes in subjective value processing and loss aversion, as mixed gambles are commonly used to assess loss aversion in decision-making under risk. Thus, we compared behavioral and neural decision-making parameters in the threat versus safe context. Furthermore, we were interested whether anticipatory anxiety would lead to connectivity changes in a putative neural valuation network and how these potential changes would relate to changes in behavioral preferences. We therefore employed psychophysiological interaction (PPI) analysis (Friston et al., 1997) to investigate changes in neural connectivity due to anticipatory anxiety.

The analysis of behavioral and neurofunctional data revealed the following picture: Electrophysiological, neurofunctional and reaction time data provided clear evidence for the successful induction of anticipatory anxiety in our participants. We found increased SCRs during threat relative to safe blocks, which serves as evidence for increased emotional arousal during threat phases relative to safe phases. Furthermore, we found peak SCRs to actual strong electrical stimulation, which increased significantly compared to SCRs to weak electrical stimulation. In line with this, the efficacy of the TOS paradigm was also reflected on the neural level as activation of a putative "pain matrix" (Casey, 1999; Tracey & Mantyh, 2007) was observed in response to strong relative to weak electrical stimulation. In addition, in the threat condition participants decided faster which option to take than in the safe condition. Although observable behavior (i.e., choices made by subjects) and decision-making parameters such as the loss aversion coefficient, λ, and prospective utility remained stable across affective contexts, we found a clear change in neural valuation signals. First, during threat trials, the ventromedial prefrontal cortex [VMPFC] and the ventral striatum [VS] showed significant reductions in coding of expected subjective value (ESV, the neural correlate of prospective utility) of the risky option. Second, functional coupling between the brain's valuation network comprising regions such as VMPFC and VS decreased during incidental anxiety. Third, while choices could be predicted from activity in VMPFC during safe trials, prediction of choices from VMPFC activity was not possible during threat trials. Fourth, we observed reduction in VMPFC and VS baseline

activity during threat relative to safe trials. In contrast to the activation pattern of VMPFC and VS, the anterior insula showed a clear increase in coding the negative ESV of the risky choice option during high incidental anxiety. Furthermore, insula activity predicted choices during threat, but not safe trials.

These patterns of results shows that incidental anxiety can lead to a shift in the focus of neural valuation from possible positive consequences to anticipated negative consequences of the choice options. These neural effects of incidental anxiety can evidently arise in the absence of changes in overt behavior. Our results may have important clinical implications, as they show that maladaptive anxiety as observed in pathological states can change the focus of neural valuation from hedonic value towards anticipated negative consequences. Our findings thereby suggest a possible pathway for the development of chronic desensitization of reward systems and a maladaptive focus on negative cognitions in the context of psychiatric disorders such as anxiety and affective disorders (e.g., Berghorst et al., 2013; Bogdan & Pizzagalli, 2006; Pizzagalli, Iosifescu, Hallett, Ratner, & Fava, 2008).

Furthermore, these findings are generally in line with previous studies on the distinct roles of VMPFC, VS and insula in decision-making showing positive versus negative value coding in VS and VMPFC versus insula, respectively (e.g., Kim, Shimojo, & O'Doherty, 2011). Both the diminished value sensitivity in VMPFC and decreased communication between VMPFC and striatal structures under anticipatory anxiety are consistent with previous rodent studies demonstrating impaired value encoding and reward learning under stress (Dias-Ferreira et al., 2009). Notably, the shift in value coding between affective contexts was not accompanied by changes in economic preferences. Our results therefore demonstrate that different neural value coding mechanisms can be associated with the same behavioral outcome. This finding seems particularly relevant in regard to the multiple systems theory of decision-making stating that different valuation systems guide choice depending on the properties of the decision-making situation (Rangel et al., 2008). The hypothesis of a change in decision-making system under anxiety is corroborated by previous studies (Schwabe, Tegenthoff, Höffken, & Wolf, 2012) reporting increased habitual instead of goal-directed responding and reduced value sensitivity in medial prefrontal cortex under induced (hormonal) stress. Since the VMPFC is commonly considered as a value integration region (Basten, Biele, Heekeren, & Fiebach, 2010), it may be speculated that neural valuation processes in VMPFC, indicated by the positive ESV signal in safe trials, broke down under anticipatory anxiety and compensatory mechanisms,

e.g., heuristics and increased habitual responding, were employed. In line with our findings, Gold and colleagues (2014) reported stable behaviors across affective contexts (threat versus safe condition) accompanied by compensatory neural processing, that is, increased connectivity of medial prefrontal cortex and VMPFC with amygdala under threat.

2.3 Study C: The neural circuitry of emotion-induced distortions of trust

With a similar experimental approach, a joint research project with Jan B. Engelmann, Christian C. Ruff, and Ernst Fehr sought to identify the interaction of incidental anxiety with social decision-making. We were interested in changes in interpersonal trust behavior due to incidental anxiety. Trust is defined as the willingness to take a social risk of helping or collaborating with another person despite the possibility of non-reciprocation. In behavioral economics, a *trust game* is supposed to mimic a sequential economic exchange in the absence of contract reinforcement institutions (Berg et al., 1995). In a standard behavioral trust game two players are anonymously paired and assigned the roles of "trustor" and "trustee". Both trustor and trustee are endowed with the same amount of money. At a first stage, the trustor decides which proportion of the original endowment he is willing to send to the trustee. Importantly, the amount the trustee receives is a multiple of the amount the trustor decided to give to the trustee. Upon receiving the, for instance, tripled amount sent by the trustor, it is the trustee's turn to decide to what extent he wants to reciprocate or take financial advantage of the experienced trust. Thus, the trustor faces two different types of risk: On the one hand, he faces a financial risk due to the possibility to lose money. On the other hand, the trust game comprises a social risk, as the trustee's response is not predictable to and controllable by the trustor.

In order to specify the impact of incidental anxiety on choice behavior that is specific to social decision-making, a well matched control task comprising a non-social instead of a social risk was included into the experiment. That is, in this control risk game a random mechanism (computer algorithm) rather than a social interaction partner was the source of risk. Whereas a classical trust game is a "one shot" game, that is, each player takes a single decision, an fMRI experiment requires multiple repetitions of such decisions. We thus developed an fMRI-compatible trust game where the trustor repeatedly faces social risks and completes the same number of matched non-social risks. Trustee responses used

in the fMRI experiment were collected in prior experimental sessions employing the strategy method.

We again chose a hybrid fMRI design in order to test the effect of incidental anticipatory anxiety (modeled in block-wise fashion) on trusting and non-social control choices (modeled in event-related fashion). As in the previously described study (section 2.2 and Appendix B), anticipatory anxiety was induced via the TOS manipulation and it was examined whether an anxious state would lead to changes in transfer rates from trustor to trustee, which served as the experimental measure of interpersonal trust.

It was conjectured that incidental anticipatory anxiety would influence trust-specific computations, i.e., the evaluation of the trustee's trustworthiness, and therefore lead to a reduced transfer rate in the threat condition compared to the safe (i.e., no threat) condition. These behavioral changes in trust due to incidental anxiety are hypothesized to be based on the modulation of neural responses in regions involved in social cognition, including temporoparietal junction (TPJ) and dorsomedial prefrontal cortex (DMPFC) (Mitchell, Macrae, & Banaji, 2006; Saxe & Kanwisher, 2003). Furthermore, we expected the amygdala and insula, regions known for their influence on emotional processing (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012), to be involved in the influence of incidental anxiety on trust decisions. VS and VMPFC, which have repeatedly been shown to be implicated in reward processing (Levy & Glimcher, 2011, 2012; Schultz, Dayan, & Montague, 1997; Sescousse, Caldú, Segura, & Dreher, 2013), were additional regions of major interest. Furthermore, we hypothesized that incidental anxiety would impact neural networks involved in emotion- (social) cognition interactions, leading to changed connectivity patterns between the network components. Thus, functional connectivity analyses were conducted in order to investigate the effect of incidental anticipatory anxiety on the valuation network as well as on emotion processing and regulation brain networks. We hypothesized that the regulation network might comprise dorsolateral prefrontal cortex (DLPFC), anterior cingulate cortex (ACC), and ventrolateral prefrontal cortex (VLPFC) (Buhle et al., 2013; Ochsner, Silvers, & Buhle, 2012; Shackman et al., 2011; Wager, Davidson, Hughes, Lindquist, & Ochsner, 2008).

On the behavioral level, we find that anticipatory anxiety reduces mean transfer rates, which were chosen as the behavioral index of trust, and mean reaction times in both the social and non-social decision-making context.

The neuroimaging results provide information about the neural mechanisms underlying the suppression of participants' propensity to trust under incidental anxiety.

First, comparing neural activity patterns for trust versus non-social (NS) control conditions, we found that TPJ, DMPFC and VMPFC were preferentially engaged during trust decisions in the safe condition ($Trust_{Safe}$ > NS $Control_{Safe}$). Second, incidental anticipatory anxiety led to a trust-specific breakdown of this activation pattern in the left TPJ, which was not observed for the control condition [assessed via the interaction: ($Trust_{Safe}$ > $Trust_{Threat}$) > (NS $Control_{Safe}$ > NS $Control_{Threat}$)]. Similarly, both VMPFC and VS show reduced trust-specific activity during threat. Third, functional connectivity analysis revealed significant connectivity between TPJ and amygdala during trust taking in the safe condition, whereas incidental anxiety was associated with a complete disruption of this connectivity. Neither did we find significant TPJ-amygdala during safe NS control trials, nor did incidental anxiety lead to a significant breakdown of TPJ-amygdala connectivity during those trials. Moreover, a trust-specific link between mean transfer rates and functional connectivity patterns in the safe context was found, which was not observed during the NS control task. That is, significantly stronger positive correlations between mean transfer levels and functional connectivity between TPJ and right amygdala, DMPFC, DLPFC, bilateral VLPFC, superior temporal sulcus (STS), and intraparietal sulcus in the trust compared to the control task were found. Connectivity strength of TPJ with these neural structures predicted mean trust levels during safe trials. However, incidental anxiety caused a breakdown of the association between mean trust taking and functional connectivity strength of TPJ specifically with posterior STS. Whereas TPJ connectivity with posterior STS predicted behavioral trust taking during the safe condition, no significant correlation between mean trust level and TPJ-posterior STS connectivity during threat was observed. As TPJ has repeatedly been implicated in representing and interpreting others' mental states and behaviors (Mitchell, 2009; Morishima, Schunk, Bruhin, Ruff, & Fehr, 2012; Young, Camprodon, Hauser, Pascual-Leone, & Saxe, 2010), our results suggest that enhanced connectivity between TPJ and regions important for social cognition (DMPFC, STS), emotion processing (amygdala) and cognitive control (VLPFC, DLPFC) supports individuals' trust taking. Incidental anticipatory anxiety appears to suppress these neural mechanisms by reducing the activation level in TPJ, reducing the connectivity between TPJ and amygdala, and disrupting the association between TPJ connectivity patterns and mean trust taking. Last, we identified a domain-general network of regions that show changes in choice-related neural activity. During threat compared to safe trials, we observed the suppression of neural activity in bilateral posterior DLPFC, left amygdala, posterior paracentral lobule, and in left VLPFC and VMPFC. Significant enhancement of activity

during decision-making under incidental anticipatory was found in thalamus and cerebellum.

These results provide potentially important insights into the neural circuitry that enables people to trust. The findings suggest that anticipatory anxiety weakens neural mechanisms important for the representation of social factors involved in interpersonal decision-making and transfer of these representations to emotion and cognitive control regions during decision-making. These changes eventually lead to a decrease in trust taking.

2.4 Study D: Impulsivity, sensation seeking and negative emotionality relate to distinct morphometric brain features

Together with Jan B. Engelmann, Ernst Fehr, and Christian C. Ruff, I investigated the neuroanatomical substrates of stable affective and personality traits predisposing towards interindividual differences in decision-making styles. Taking into account the numerous studies conducted on the neurobiology of those traits, we were especially interested in an improved assessment of personality in order to be able to conclude about variance in brain structure that is uniquely explained by variation in one personality feature and not confounded by other constructs. We applied a comprehensive set of personality questionnaires covering traits and characteristics identified to be associated with interindividual variation in decision-making style and were particularly interested in potential colinearity between the single personality constructs assessed. We scrutinized the battery of psychological questionnaires to factor analysis in order to directly address the possible colinearity of the personality features of interest. We then combined the resulting factor scores, which indicate the subjects' position in the personality dimension identified by the factor analysis, with subjects' anatomical brain information collected in the realm of two other studies presented in this thesis (Study B and Study C, n = 80). The factor analysis identified three factors, confirming a crucial distinction between impulsivity and risk-taking (factor 1), negative emotionality (factor 2) and sensation seeking (factor 3), which have been shown to influence daily decision-making (Manning et al., 2014; Miu, Heilman, & Houser, 2008). Following, we employed voxel-based morphometry (VBM; Ashburner,

2007) in order to identify systematic relationships between interindividual variations in brain anatomy and the personality and affective characteristics identified via FA.

Taken together, the investigation on the neuroanatomical correlates of personality characteristics leading to differences in individual decision-making revealed unique brain morphometric correlates of impulsivity, sensation seeking, and negative emotionality. The VBM analysis confirmed a substantial distinction between impulsivity and sensation seeking: Increased impulsivity and sensation seeking were uniquely associated with reduced grey matter (GM) volume in medial orbitofrontal cortex (mOFC) and anterior cingulate cortex (ACC), respectively. Moreover, negative emotionality showed a negative relationship with GM volume in dorsolateral prefrontal cortex (DLPFC) and inferior parietal cortex, which seems congruent with reported changes in fronto-parietal networks in depressive patients reported by Schutter et al. (2002). No significant positive associations between the identified personality factors and the neuroanatomical regions of interest were obtained.

In conclusion, our study elucidates the neurobiological basis of personality traits that modify individual decision-making and suggests that nonclinical individual differences in brain morphometry may constitute a neurobiological predisposition for impulsivity, sensation seeking, and negative emotionality, which may modify decision-making dependent on contextual and cognitive demands (Morishima et al., 2012). More specifically, by simultaneously assessing several of these personality traits and accounting for their shared variance in the analysis, we were able to distinguish two commonly confounded concepts, impulsivity, and sensation seeking, on the behavioral and neural level. These findings also suggest that neurobiological information may be usefully employed in psychiatric research, since aberrant personality features and their modifying influence on decision-making are a common symptom of psychiatric disorders. This distinction could be especially important in the domain of diagnostics and treatment of addiction. It has been shown that impulsivity can be a risk factor for the development of a stimulant dependency, whereas increased sensation seeking seemed to be a consequence of the stimulant addiction (Ersche et al., 2010). Our investigation thus informs psychiatric research and might lead to new impulses for the diagnostics of clinical disorders. Furthermore, the current study suggests that research on the relation between brain structure and personality traits clearly benefits from an approach that explicitly measures the colinearity and conceptual overlap of competing personality constructs.

3 General Discussion

In the following sections, I will summarize the results of the conducted studies and draw conclusions regarding the effects of emotional factors on decision-making derived from the presented studies. Furthermore, I will formulate questions arising from this dissertation and directions for future research.

The presented studies are intended to contribute to the rising field of research on the modulatory influence of emotion on decision-making. Different methodological approaches such as fMRI, VBM and TMS were employed in order to gain a comprehensive overview of the modulation of human decision-making by different emotion-related factors on the behavioral, neuroanatomical and neurofunctional level. The conducted studies add to the knowledge of the modulatory role of emotion on decision-making and emphasize the highly task- and context-dependent effects of emotion on decision-making. That is, the interaction between emotional factors and decision-making appears to be highly dependent on the quality of the emotional response elicited and the characteristics of the decision-making situation at hand (see also Lerner et al., 2015; Phelps et al., 2014).

The conducted TMS study (section 2.1 and Appendix A) was intended to identify the role of the DLPFC in resistance to risky choice framing effects. The framing effect has been described as an emotional bias in decision-making (Cassotti et al., 2012; De Martino et al., 2006; Xu et al., 2013) and further evidence for its emotional basis has been confirmed in a recent study (Stanton, Reeck, Huettel, & LaBar, 2014). Although previous neuroimaging studies (De Martino et al., 2006; Gonzalez et al., 2005) reported correlative evidence for the role of DLPFC in framing, we did not find changes in choice and preference as a function of temporary unilateral DLPFC activity disruption. Thus, we cannot confirm the functional relevance of unilateral DLPFC in inhibition of framing susceptibility. I assume that DLPFC may implement computations that are correlated with neural processes relevant for framing resistance, but these computations do not contribute to framing resistance itself. Although both DLPFC and ACC are key structures of a neural network of cognitive control (e.g., Mansouri, Tanaka, & Buckley, 2009), the ACC, a structure known for its involvement in conflict monitoring and emotional processing (Botvinick, Cohen, & Carter, 2004; Bush, Luu, & Posner, 2000; Shackman et al., 2011), might be the fundamental neural component in the implementation of framing resistance, i.e., the suppression of affective biases elicited

by decision frames and emotional reactions to these frames. Methodological modifications such as bilateral TMS of DLPFC, dual site TMS of DLPFC and ACC and bilateral or dual site tDCS routines should be employed to enable strong conclusions about the causal mechanisms leading to framing effects in risky choice. Our result constricts the possible neurobiological basis of resistance to the framing effect by showing that at least unilateral DLPFC disruption is not sufficient to change the influence of integral emotion on decision-making. As our findings do not allow a conclusion about the role of DLPFC, future studies should employ methodological changes as described above to reach stronger conclusions regarding the correlational versus causal involvement of DLPFC in framing.

Considering the effects of incidental anxiety on non-social (section 2.2 and Appendix B) and social decision-making (section 2.3 and Appendix C), a differentiated picture appears, pointing to task- and social context-dependent effects of anticipatory anxiety on the behavioral and neural correlates of human economic preference. That is, the study on the effects of anticipatory anxiety on behavioral and neural economic value processing in mixed gambles, which did not contain an explicit social dimension as choices taken did not imply consequences for other nor were they dependent on other people's choices, revealed decreased neural value coding accompanied by stable behavioral preferences under threat, i.e., task-unrelated anticipatory anxiety. This observation might be interpreted as reduced hedonic capacity, i.e., decreased reward responsiveness, in a state of anxiety (Berghorst et al., 2013; Bogdan & Pizzagalli, 2006). Although we found changes in neural valuation, no changes in behavioral preferences due to incidental anticipatory anxiety were found. In line with findings by Gold and colleagues (2014), who report stable goal-directed behavior combined with changed neural activity and connectivity patterns under threat relative to safety using experimental procedures similar to ours, our results may bear clinical implications. Specifically, although neural correlates of valuation and expected subjective value in persons suffering from psychiatric conditions can change, the persons concerned can show non-deviant preferences at early (i.e., prodromal) stages. Still, their hedonic capacity could be impacted, potentially reflected in decreased neural reward responsiveness (however, see Bogdan & Pizzagalli (2006) reporting *observable* behavioral changes in reward responsiveness due to acute anxiety). Reduced hedonic capacity is regularly observed in several psychiatric disorders such as depression, which is often characterized by pathological anhedonia, the reduced ability to experience pleasure (Treadway & Zald, 2011). During the progressive course of a psychiatric disease such as anxiety or depression, neural valuation processes can completely deteriorate to the point

where observable behavior also breaks down (for review, see Sharp, Monterosso, & Montague, 2012) and we can observe neural *and* behavioral changes due to maladaptive emotional and psychiatric conditions. A still open question is how exactly the behavioral stability is perpetuated during these prodromal stages. Future studies should investigate whether and under which conditions the effect of anxiety on behavior is compensated for by functional connectivity adaptation as suggested by our study and Gold et al. (2014). Moreover, it should be explored which other potential mechanisms – apart from changes in functional connectivity – may lead to the dissociation between behavioral and neural valuation in anxious states. It needs to be systematically investigated how blunted neural reward responsiveness is translated into behavior. Might we, for example, expect decreased or increased risk-taking in people suffering from anxiety disorders? Inconsistent results have been reported. Giorgetta et al. (2012) and Grecucci et al. (2013) observed decreased risk-taking in anxious subjects, whereas Kashdan and colleagues (2006) found increased risk-taking in socially anxious subjects. Montoya et al. (2014) found increased risk-taking after cortisol administration, a procedure mimicking stress and anxious responses. The authors hypothesized that this increase in risk-taking might be attributed to the organism trying to compensate the hypoactive state of the reward system under stress by increased risk-taking.

In contrast to the study on the effects of incidental anticipatory anxiety on non-social decision-making, we did find changes in interpersonal trust and associated neural correlates due to incidental anticipatory anxiety (section 2.3 and Appendix C). Specifically, we observed that decreases in trusting behavior due to anxiety correlated with changes in social cognition-specific and valuation brain networks. On the behavioral level, incidental anxiety also decreased transfer rates in a matched non-social control risk game to a similar extent as in the trust game. However, the neural circuitry underlying these behavioral changes due to anxiety significantly differed between the trust and control condition. During trust choices, left TPJ, a region known for its involvement in social cognition (e.g., Samson, Apperly, Chiavarino, & Humphreys, 2004), was active in the safe context, but broke down under threat. During risky decisions, no significant activation of TPJ in the safe or threat context was observed. Furthermore, a neural network typically involved in social cognition (e.g., Samson et al., 2004; Saxe, 2006) consisting of left TPJ, amygdala, DMPFC and VMPFC showed significantly greater connectivity during decision-making in trust relative to control trials in the safe condition. Under threat, we observed a trust-specific breakdown of these connectivity patterns. In addition, brain-behavior relationships, that is,

transfer rates correlating with connectivity strength between TPJ and posterior STS, observed under safe conditions broke down under threat. The posterior STS is a region essential for inferences concerning other people's intentions (Saxe, 2006; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004), a core component of social cognition (for review, see Amodio & Frith, 2006). We might speculate that acute anxiety impacts the neural representation of others' intentional actions and the adequate interpretation of social context, thus leading to changes in social interaction behavior. The results from this study inform us about the fundamental role of incidental emotions in social decision-making and suggest that the integration of incidental emotion in economic models of decision-making could improve the predictive power of these models (Lerner et al., 2015; Loewenstein & Lerner, 2003).

Furthermore, comparing the results of Studies B and C investigating the effect of incidental anticipatory anxiety on non-social versus social decision-making, respectively, I conclude that the carryover effects of incidental anxiety on choice and judgment appear to be highly specific, depending on both the quality of the emotion induced and the decision-making task characteristics (for review, see Phelps et al., 2014). While neural changes (i.e., decreased neural sensitivity to economic value) accompanied by stable behavioral preferences were found in (non-social) mixed gambles under incidental anxiety (section 2.2 and Appendix B), we observed both behavioral and neural changes in the trust game and control risk game under threat (section 2.3 and Appendix C). Taking into account the experimental designs employed in the two studies, we might hypothesize that the impact of emotion on decision-making is dependent on task complexity. This conjecture is based on the fact that the trust game and adapted risk game are characterized by increased complexity compared to the mixed gambles, which comprise equiprobability lotteries that are relatively simple from a cognitive point of view. Hence, we might speculate that, on the one hand, the effects of emotion on decision-making are highly dependent on the cognitive task load (Berggren, Richards, Taylor, & Derakshan, 2013; Pessoa, 2009; Pessoa, Padmala, Kenzer, & Bauer, 2012). On the other hand, we should ask for specific interactions of emotion with social versus non-social decision-making contexts, as neural networks impacted by these distinct contexts have been suggested to differ (V. Lee & Harris, 2013; Ruff & Fehr, 2014), which is also demonstrated in our study. In line with this, early behavioral studies demonstrated stronger effects of mood on social decision-making when tasks were demanding and complex (Forgas, 1995). Thus, stronger emotional effects on decision-making seem to occur for more elaborate interactions in social contexts. Harlé et al. (2012)

hypothesized that socially more complex choice situations may be more susceptible to emotional influences than non-social and less complex tasks. That is, those kinds of decision-making situations draw on more behavioral cognitive resources and underlying neural processing capacities and hence should be prone to interference by parallel processes. Considering the results of the current studies, a complex picture, which needs differentiation, appears. On the one hand, incidental anxiety did not interfere with observable behavior in a simple risky decision-making task (section 2.2 and Study B), which is to some degree consistent with the hypothesis of Harlé et al. (2012) that affect infusion onto decision-making processes is more likely for cognitively complex tasks. On the other hand, in Study C, we found that social and non-social decision-making are equally susceptible to incidental emotional interference. That is, transfer rates in trust and risk game decreased to similar degrees in the threat relative to safe context. Importantly, the two tasks were well matched regarding cognitive complexity. Not the social or non-social characteristic of the decision-making situation determined the strength of the emotion-cognition interaction; rather the complexity of the task seemed to make our subjects' decision-making more or less susceptible to emotional influence.

Furthermore, an important distinction between revealed preferences and associated neural correlates was revealed by Study B: We demonstrated that the brain's value coding changes between safe and threat emotional contexts, with a focus on negative value coding during anxiety, whereas observed behavior (i.e., choices) were constant across emotional contexts. This finding is generally congruent with multiple systems theories of decision-making stating that different valuation systems can guide choice depending on the properties of the decision-making situation (Rangel et al., 2008). Changes in neural valuation processes due to negative affective states are also reported by Schwabe et al. (2012). They reported that hormonally induced stress can lead to a change from goal-directed to habitual decision control and reduced value sensitivity in medial prefrontal cortex. During anticipatory anxiety, the same behavioral choices may have been based on goal-directed negative value computations in the insula (rather than positive value computations in the VMPFC and VS), possibly in conjunction with compensatory mechanisms such as heuristics and increased habitual responding. In either case, the reduction of value coding in the VMPFC and VS during anxiety is suggested to indicate a decrease in the positive value of the available choice options.

In addition, our data add to the interesting discussion on the instantiation of social versus non-social valuation in the brain (Ruff & Fehr, 2014). We found distinct brain

networks to be involved in social versus non-social valuation processes (Study C). Whereas social valuation was modulated by social cognition-related brain structures (e.g., left TPJ, pSTS) and general valuation brain regions (VMPFC, DLPFC, VLPFC), non-social choices were computed in general valuation brain regions without prominent computation of value in social cognition-specific brain regions, which we instead observed for trust decisions.

Although our data do not allow a firm conclusion on the neural architecture of social versus non-social value processing, based on these results we may conceptualize the instantiation of social valuation in the brain according to the *extended common currency* schema (Ruff & Fehr, 2014). According to this schema, identical neural processes determine the motivational relevance for social and non-social decision objects. Specifically, a general neural valuation network is assumed to integrate input from brain regions relevant for social and non-social choice in order to compute specific social or non-social value signals, respectively, entering the neural utility computation. It is predicted that social and non-social neural value signals are processed in the same valuation brain regions. Depending on the actual "value domain", functional connectivity patterns with distinct brain regions or networks (e.g., social cognition circuitry during social choice) will vary. Alternatively, the authors propose the *social-valuation-specific* schema, which presumes that social values are processed in specifically dedicated neural networks. According to this schema, different neurons are proposed to process social versus non-social factors. Social and non-social neural value signals are implemented via distinct spatial patterns of brain activity. Importantly, the underlying computations in social and non-social value processing follow similar principles.

Our study highlights social cognition-relevant brain structures as neural correlates of the behaviorally observed breakdown of trust during anxiety in combination with general valuation specific brain regions during trust decisions. The observed domain-general effects of incidental anxiety on decision-making and valuation relevant brain regions such as VMPFC, VLPFC, and DLPFC might also support the extended common currency schema. However, we cannot finally conclude on the mechanisms underlying the observed changes: Does the affective manipulation impact a "general value-representation circuit" and – depending on choice domain – specific functional connectivity patterns? Alternatively, does anticipatory anxiety impact on a distinct "social value-representation circuit" versus a "non-social value-representation circuit" in social versus non-social choice contexts, respectively? Future studies specifically designed to disentangle social versus non-social valuation and decision-making in the brain may employ the above described taxonomy.

Furthermore, distinct effects of emotional factors on social and non-social decision-making should be considered in order to gain a systematic understanding of the complex interactions between emotions, decision-making and choice context.

Furthermore, we should take notice of the different samples used in the two studies. While a male sample was used for the non-social decision-making task, a mixed gender sample was used for the social decision-making task. We might speculate that gender differences are important in determining the effect of emotion on decision-making (Fehr-Duda, Epper, Bruhin, & Schubert, 2011; Lighthall, Mather, & Gorlick, 2009). Fehr-Duda et al. (2011), for instance, reported that incidental good mood was significantly correlated with probability weighting in women. That is, females in a better than normal mood tend to weight probabilities relatively more optimistically. Men, though, appear immunized against effects of incidental mood on decision-making by applying a mechanical decision criterion such as the maximization of expected value. We might tentatively speculate about "spontaneous" decision-making instead of employing mechanistic decision rules in the trust game study. Therefore, future studies may tackle the issues implied by sample characteristics. At this point, we should be hesitant concerning the generalizability of our results as effects might be significantly driven by our sample characteristics.

In conclusion, the conducted fMRI studies offer empirical evidence for the diverse and highly specific influences of emotional factors on decision-making. I found that dynamic emotional factors systematically modulate neurofunctional correlates of human decision-making. Our results thus support the call for the integration of incidental emotional factors into economic models of decision-making. The finding of decreased trusting behavior in a state of incidental anticipatory anxiety in the fMRI study is congruent with psychological perspectives on emotional influences on decision-making. For example, the ATF (Lerner & Keltner, 2000, 2001) predicts decreased trust in an anxious state due to an appraisal of an uncertain existential threat, leading to risk-avoidant behavior in order to reduce this uncertainty. Future research needs to decide about an adequate model explaining the manifold and diverse effects of emotion on decision-making. At this point we can state that valence-grounded approaches are a parsimonious, but too simple way to explain the effects of emotion on decision-making. Previous studies presented behavioral data that was more consistent with the ATF by Lerner and Keltner (2001). Neuroimaging data can enrich this research and help reveal the manifold, but consistent effects of emotion on choice and preference. As stated above, the investigation of the neural organization of social versus non-social value processing, especially in combination with emotional influences, is another

very interesting topic. Taking Ruff and Fehr's schemas (2014) as a conceptual framework seems appropriate in order to examine the neural representation of social concepts and understand healthy interactive behavior as well as deviations and maladaptive behavior in psychopathology. To the best of our knowledge, the sparse number of studies on the interaction of emotion and decision-making used correlative measures. In order to identify causal mechanisms underlying these interactions and distinguish them from purely correlated activity not causally linked to the behavior in question, future studies need to apply combined brain stimulation and neuroimaging techniques such as concurrent TMS-fMRI in order to observe the effects of the interruption of putative "social networks". However, this approach is limited by the subcortical localization of brain structures fundamentally involved in emotion and reward processing, so that TMS application will be restricted to cortical regions such as DLPFC and to the observation of network effects on subcortical structures (e.g., striatum or amygdala) via fMRI.

Finally, our interest in the neurobiology of personality and affective traits relevant for decision-making led to the design of a VBM study in order to investigate the neuroanatomical correlates of these traits. This study elucidates the neurobiological basis of personality traits that modify individual decision-making. By simultaneously assessing several of these personality and affective traits and accounting for their shared variance in the analysis, we were able to distinguish two commonly confounded concepts, impulsivity and sensation seeking, on the behavioral and neural level. While we found that higher impulsivity and risk-taking were associated with smaller mOFC GM volume, a negative relationship between sensation seeking and ACC GM volume was observed. This distinction might be especially important in the diagnostics and treatment of psychiatric disorders such as addiction. That is, it has been shown that impulsivity could be a risk factor for the development of a stimulant dependency, whereas increased sensation seeking seemed to be a consequence of the stimulant addiction (Ersche et al., 2010). In addition, we identify the neuroanatomical correlate of negative emotionality, that is, higher negative emotionality scores were correlated with smaller regional DLPFC and inferior parietal GM volume. Negative emotionality has repeatedly been shown to be associated with psychiatric disorders such as depression, alcohol dependence and borderline personality disorder (Boschloo et al., 2013; Ruocco, Amirthavasagam, Choi-Kain, & McMain, 2013). Hence, our investigation informs psychiatric research and might lead to impulses for the diagnostics of clinical disorders by supporting the identification of endophenotypes of psychiatric disorders, therefore advancing dimensional psychiatric diagnostics. Of course,

the generalization of our results to psychiatric cases needs to be taken with caution, as our subjects were mentally healthy persons from a student pool. However, as healthy and pathological personality traits build a continuum (e.g., Grupe & Nitschke, 2013), it is essential to understand neurobiological mechanisms underlying those traits in healthy samples to gain a systematic understanding of "normal" neuroanatomy-personality associations and deviations in psychopathology with the goal to improve psychiatric diagnostics and treatment. At a more general level, our investigation suggests that research on the relation between brain structure and stable traits clearly benefits from an approach that explicitly measures the colinearity and conceptual overlap of competing personality constructs.

To put it in a nutshell, we found diverse evidence for the neurofunctional and neuroanatomical representation of incidental emotional effects on decision-making, contradicting the economic dogma of the inferior role of incidental emotion in decision-making. In order to extend the knowledge presented in this thesis and systematize affective influences on choice and preferences in the behavioral and neural domain, I propose to use the emotion-imbued choice (EIC) model by Lerner et al. (2015), a general framework of emotional influences on decision-making. This model advances the integration of both rational choice theory and findings on integral and incidental emotional influences on decision-making in order to develop empirically valid and naturalistic economic models of human decision-making. The model comprises components consistent with normative models of decision-making. That is, each option's expected utility of a choice at hand (including the assessment of expected emotions associated with each outcome) is assessed separately. The resulting utilities are then combined with choice characteristics (e.g., probabilities, time delays) as well as decision maker characteristics (e.g., risk preferences) to form an overall evaluation of each option. Importantly, emotion enters the model in two ways: Each outcome's utility is also assessed in regard to one's expected emotional response to that outcome (see above). These predicted emotions are compatible with rational choice theory and contribute to utility calculation as rational factors. However, the second type of emotion, namely emotions that are felt at the time of decision-making, is incompatible with rational choice theory. It is assumed that there are distinct sources of these current emotions: (1) Characteristics of the choice options can influence the decision-maker's current emotional state. (2) Expected emotions may not only directly impact outcome evaluation, but also impact the current emotional state. (3) Evaluation of the choice situation may elicit emotional responses such as frustration, which leads to changes

in current emotional state. (4) Incidental emotions such as baseline or dispositional anxiety, mood or weather impact the decision maker's current feelings in addition to the above described integral factors (1-3). These current feelings, in turn, influence the evaluation of the choice inputs. Depending on the emotional state, different computations dominate the decision-making process (e.g., heuristic versus analytic processing activated motivational goals, or choice dimensions focused on). The EIC model is based on the central tenets of the ATF, and it seems well suited to guide future research on the neurobiological and behavioral mechanisms underlying emotional influences on choice and judgment.

Future research should employ this general framework in order to develop an adequate model, which synthesizes assumptions of rational choice theory and emotional factors to explain human decision-making. I hypothesize that this integration will improve the predictive power of economic models of choice and increase their applicability in contexts such as politics, consumer choice, and domains like psychiatric research.

Together, the findings presented in this thesis provide empirical evidence for the modulation of decision-making by incidental and integral emotional sources on the behavioral and neurobiological level calling for the consistent integration of emotion into economic theories of decision-making.

References

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277. doi:10.1038/nrn1884

Andrade, E., & Ariely, D. (2009). The enduring impact of transient emotions on decision making. *Organizational Behavior and Human Decision Processes*, *109*(1), 1–8.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, *38*(1), 95–113. doi:10.1016/j.neuroimage.2007.07.007

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412–427. doi:10.1016/j.neuroimage.2013.02.063

Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21767–21772. doi:10.1073/pnas.0908104107

Berggren, N., Richards, A., Taylor, J., & Derakshan, N. (2013). Affective attention under cognitive load: reduced emotional biases but emergent anxiety-related costs to inhibitory control. *Frontiers in Human Neuroscience*, *7*. doi:10.3389/fnhum.2013.00188

Berghorst, L. H., Bogdan, R., Frank, M. J., & Pizzagalli, D. A. (2013). Acute stress selectively reduces reward sensitivity. *Frontiers in Human Neuroscience*, *7*, 133. doi:10.3389/fnhum.2013.00133

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, *10*(1), 122–142. doi:10.1006/game.1995.1027

Bevilacqua, L., & Goldman, D. (2013). Genetics of impulsive behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1615). doi:10.1098/rstb.2012.0380

Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion*, *24*(4), 561–595. doi:10.1080/02699930903132496

Bodenhausen, G. V., Sheppard, L. A., & Kramer, G. P. (1994). Negative affect and social judgment: The differential impact of anger and sadness. *European Journal of Social Psychology*, *24*(1), 45–62. doi:10.1002/ejsp.2420240104

Bogdan, R., & Pizzagalli, D. A. (2006). Acute Stress Reduces Reward Responsiveness: Implications for Depression. *Biological Psychiatry*, *60*(10), 1147–1154. doi:10.1016/j.biopsych.2006.03.037

Bonini, N., Hadjichristidis, C., Mazzocco, K., Demattè, M. L., Zampini, M., Sbarbati, A., & Magon, S. (2011). Pecunia olet: the role of incidental disgust in the ultimatum game. *Emotion (Washington, D.C.)*, *11*(4), 965–969. doi:10.1037/a0022820

Boschloo, L., Vogelzangs, N., van den Brink, W., Smit, J. H., Beekman, A. T. F., & Penninx, B. W. J. H. (2013). The role of negative emotionality and impulsivity in depressive/anxiety disorders and alcohol dependence. *Psychological Medicine*, *43*(6), 1241–1253. doi:10.1017/S0033291712002152

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, *8*(12), 539–546. doi:10.1016/j.tics.2004.10.003

Bouchard Jr, T. J., & Loehlin, J. C. (2001). Genes, Evolution, and Personality. *Behavior Genetics*, *31*(3), 243–273. doi:10.1023/A:1012294324713

Buhle, J. T., Silvers, J. A., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., … Ochsner, K. N. (2013). Cognitive Reappraisal of Emotion: A Meta-Analysis of Human Neuroimaging Studies. *Cerebral Cortex (New York, N.Y.: 1991)*. doi:10.1093/cercor/bht154

Bush, G., Luu, P., & Posner, M.I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*(6), 215–222.

Camerer, C. (2005). Three Cheers – Psychological, Theoretical, Empirical – For Loss Aversion. *Journal of Marketing Research*, *42*(2), 129–133. doi:10.2307/30164010

Canli, T., Zhao, Z., Desmond, J. E., Kang, E., Gross, J., & Gabrieli, J. D. (2001). An fMRI study of personality influences on brain reactivity to emotional stimuli. *Behavioral Neuroscience*, *115*(1), 33–42.

Caplin, A., & Leahy, J. (2001). Psychological Expected Utility Theory and Anticipatory Feelings. *Quarterly Journal of Economics*, *116*(1), 55–79. doi:10.1162/003355301556347

Casey, K. L. (1999). Forebrain mechanisms of nociception and pain: Analysis through imaging. *Proceedings of the National Academy of Sciences*, *96*(14), 7668–7674. doi:10.1073/pnas.96.14.7668

Cassotti, M., Habib, M., Poirel, N., Aïte, A., Houdé, O., & Moutier, S. (2012). Positive emotional context eliminates the framing effect in decision-making. *Emotion (Washington, D.C.)*, *12*(5), 926–931. doi:10.1037/a0026788

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. *The Journal of Neuroscience*, *29*(39), 12315–12320. doi:10.1523/JNEUROSCI.2575-09.2009

Clark, L., Li, R., Wright, C. M., Rome, F., Fairchild, G., Dunn, B. D., & Aitken, M. R. F. (2012). Risk-avoidant decision making increased by threat of electric shock. *Psychophysiology*, *49*(10), 1436–1443. doi:10.1111/j.1469-8986.2012.01454.x

Cloninger, C. R. (1999). *The Temperament and Character Inventory—Revised*. St. Louis, MO: Center for Psychobiology of Personality, Washington University.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Coutlee, C. G., & Huettel, S. A. (2012). The functional neuroanatomy of decision making: prefrontal control of thought and action. *Brain Research*, *1428*, 3–12. doi:10.1016/j.brainres.2011.05.053

Davis, C., Patte, K., Tweed, S., & Curtis, C. (2007). Personality traits associated with decision-making deficits. *Personality and Individual Differences*, *42*(2), 279–290. doi:10.1016/j.paid.2006.07.006

De Martino, B., Camerer, C. F., & Adolphs, R. (2010). Amygdala damage eliminates monetary loss aversion. *Proceedings of the National Academy of Sciences*, *107*(8), 3788–3792. doi:10.1073/pnas.0910230107

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science (New York, N.Y.)*, *313*(5787), 684–687. doi:10.1126/science.1128356

Denny, B. T., Ochsner, K. N., Weber, J., & Wager, T. D. (2013). Anticipatory brain activity predicts the success or failure of subsequent emotion regulation. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nss148

DeYoung, C. G., & Gray, J. R. (2009). Personality Neuroscience: Explaining Individual Differences in Affect, Behavior, and Cognition. In P. J. Corr & G. Matthews (Eds.), *Cambridge handbook of personality*. New York: Cambridge University Press.

DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., & Gray, J. R. (2010). Testing predictions from personality neuroscience. Brain structure and the big five. *Psychological Science*, *21*(6), 820–828. doi:10.1177/0956797610370159

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., … Sousa, N. (2009). Chronic Stress Causes Frontostriatal Reorganization and Affects Decision-Making. *Science*, *325*(5940), 621–625. doi:10.1126/science.1171203

Domschke, K., & Deckert, J. (2012). Genetics of anxiety disorders - status quo and quo vadis. *Current Pharmaceutical Design*, *18*(35), 5691–5698.

Druckman, J. N., & McDermott, R. (2008). Emotion and the Framing of Risky Choice. *Political Behavior*, *30*(3), 297–321. doi:10.1007/s11109-008-9056-y

Ersche, K. D., Turton, A. J., Pradhan, S., Bullmore, E. T., & Robbins, T. W. (2010). Drug addiction endophenotypes: impulsive versus sensation-seeking personality traits. *Biological Psychiatry*, *68*(8), 770–773. doi:10.1016/j.biopsych.2010.06.015

Fecteau, S., Knoch, D., Fregni, F., Sultani, N., Boggio, P., & Pascual-Leone, A. (2007). Diminishing risk-taking behavior by modulating activity in the prefrontal cortex: a direct current stimulation study. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(46), 12500–12505. doi:10.1523/JNEUROSCI.3283-07.2007

Fehr-Duda, H., Epper, T., Bruhin, A., & Schubert, R. (2011). Risk and rationality: The effects of mood and decision rules on probability weighting. *Journal of Economic Behavior & Organization*, *78*(1–2), 14–24. doi:10.1016/j.jebo.2010.12.004

Figner, B., Knoch, D., Johnson, E. J., Krosch, A. R., Lisanby, S. H., Fehr, E., & Weber, E. U. (2010). Lateral prefrontal cortex and self-control in intertemporal choice. *Nature Neuroscience*, *13*(5), 538–539. doi:10.1038/nn.2516

Fischer, A. H., & Manstead, A. S. R. (2009). Social functions of emotion. In M. Lewis, J. Haviland-Jones, & L. Feldman-Barrett (Eds.), *Handbook of Emotions.* (3rd ed., pp. 456–468). New York: Guilford.

Fox, C. R., & Poldrack, R. A. (2009). Prospect Theory and the Brain. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision Making and the Brain*. New York: Academic Press.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and Modulatory Interactions in Neuroimaging. *NeuroImage*, *6*(3), 218–229. doi:10.1006/nimg.1997.0291

Giorgetta, C., Grecucci, A., Zuanon, S., Perini, L., Balestrieri, M., Bonini, N., … Brambilla, P. (2012). Reduced risk-taking behavior as a trait feature of anxiety. *Emotion (Washington, D.C.)*, *12*(6), 1373–1383. doi:10.1037/a0029119

Gold, A. L., Morey, R. A., & McCarthy, G. (2014). Amygdala-Prefrontal Cortex Functional Connectivity During Threat-Induced Anxiety and Goal Distraction. *Biological Psychiatry*. doi:10.1016/j.biopsych.2014.03.030

Gonzalez, C., Dana, J., Koshino, H., & Just, M. (2005). The framing effect and risky decisions: Examining cognitive functions with fMRI. *Journal of Economic Psychology*, *26*(1), 1–20. doi:10.1016/j.joep.2004.08.004

Grecucci, A., Giorgetta, C., Brambilla, P., Zuanon, S., Perini, L., Balestrieri, M., Sanfey, A. G. (2013). Anxious ultimatums: how anxiety disorders affect socioeconomic behaviour. *Cognition & Emotion*, *27*(2), 230–244. doi:10.1080/02699931.2012.698982

Grupe, D. W., & Nitschke, J. B. (2011). Uncertainty is associated with biased expectancies and heightened responses to aversion. *Emotion (Washington, D.C.)*, *11*(2), 413–424. doi:10.1037/a0022583

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience*, *14*(7), 488–501. doi:10.1038/nrn3524

Han, S., Lerner, J. S., & Keltner, D. (2007). Feelings and consumer decision making: The appraisal-tendency framework. *Journal of Consumer Psychology*, *17*, 158–168.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science (New York, N.Y.)*, *324*(5927), 646–648. doi:10.1126/science.1168450

Harlé, K. M., Chang, L. J., van 't Wout, M., & Sanfey, A. G. (2012). The neural mechanisms of affect infusion in social economic decision-making: a mediating role of the anterior insula. *NeuroImage*, *61*(1), 32–40. doi:10.1016/j.neuroimage.2012.02.027

Harlé, K. M., & Sanfey, A. G. (2007). Incidental sadness biases social economic decisions in the Ultimatum Game. *Emotion (Washington, D.C.)*, *7*(4), 876–881. doi:10.1037/1528-3542.7.4.876

Harlé, K. M., & Sanfey, A. G. (2010). Effects of approach and withdrawal motivation on interactive economic decisions. *Cognition & Emotion*, *24*(8), 1456–1465. doi:10.1080/02699930903510220

Hartley, C. A., & Phelps, E. A. (2012). Anxiety and decision-making. *Biological Psychiatry*, *72*(2), 113–118. doi:10.1016/j.biopsych.2011.12.027

Hayward, G., Mehta, M. A., Harmer, C., Spinks, T. J., Grasby, P. M., & Goodwin, G. M. (2007). Exploring the physiological effects of double-cone coil TMS over the medial frontal cortex on the anterior cingulate cortex: an H2(15)O PET study. *The European Journal of Neuroscience*, *25*(7), 2224–2233. doi:10.1111/j.1460-9568.2007.05430.x

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science (New York, N.Y.), 310*(5754), 1680–1683. doi:10.1126/science.1115327

Hu, X., Erb, M., Ackermann, H., Martin, J. A., Grodd, W., & Reiterer, S. M. (2011). Voxel-based morphometry studies of personality: issue of statistical model specification–effect of nuisance covariates. *NeuroImage*, *54*(3), 1994–2005. doi:10.1016/j.neuroimage.2010.10.024

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. *Biology of Mood & Anxiety Disorders*, *3*(1), 12. doi:10.1186/2045-5380-3-12

Inbar, Y., Pizarro, D. A., Knobe, J., & Bloom, P. (2009). Disgust sensitivity predicts intuitive disapproval of gays. *Emotion (Washington, D.C.)*, *9*(3), 435–439. doi:10.1037/a0015960

Johnson, E. J., & Tversky, A. (1983). Affect, Generalization, and the Perception of Risk. *Journal of Personality and Social Psychology*, *45*(1), 20–31.

Joseph, J. E., Liu, X., Jiang, Y., Lynam, D., & Kelly, T. H. (2009). Neural correlates of emotional reactivity in sensation seeking. *Psychological Science*, *20*(2), 215–223. doi:10.1111/j.1467-9280.2009.02283.x

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263. doi:10.2307/1914185

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*, *12*(4), 231–242. doi:10.1038/nrn3000

Kashdan, T. B., Collins, R. L., & Elhai, J. D. (2006). Social Anxiety and Positive Outcome Expectancies on Risk-Taking Behaviors. *Cognitive Therapy and Research*, *30*(6), 749–761. doi:10.1007/s10608-006-9017-x

Keltner, D., & Lerner, J. S. (2010). Emotion. In *Handbook of Social Psychology* (pp. 312–347). John Wiley & Sons, Inc.

Kleef, G. A. V. (2009). How Emotions Regulate Social Life The Emotions as Social Information (EASI) Model. *Current Directions in Psychological Science*, *18*(3), 184–188. doi:10.1111/j.1467-8721.2009.01633.x

Kim, H., Shimojo, S., & O'Doherty, J. P. (2011). Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, *21*(4), 769–776. doi:10.1093/cercor/bhq145

Kishida, K. T., King-Casas, B., & Montague, P. R. (2010). Neuroeconomic Approaches to Mental Disorders. *Neuron*, *67*(4), 543–554. doi:10.1016/j.neuron.2010.07.021

Kishida, K. T., & Montague, P. R. (2013). Economic probes of mental function and the extraction of computational phenotypes. *Journal of Economic Behavior & Organization*, *94*(100), 234–241. doi:10.1016/j.jebo.2013.07.009

Knoch, D., Gianotti, L. R. R., Pascual-Leone, A., Treyer, V., Regard, M., Hohmann, M., & Brugger, P. (2006). Disruption of right prefrontal cortex by low-frequency repetitive transcranial magnetic stimulation induces risk-taking behavior. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(24), 6469–6472. doi:10.1523/JNEUROSCI.0804-06.2006

Kühberger, A., & Tanner, C. (2010). Risky choice framing: Task versions and a comparison of prospect theory and fuzzy-trace theory. *Journal of Behavioral Decision Making*, *23*(3), 314–329. doi:10.1002/bdm.656

Kumari, V., Ffytche, D. H., Williams, S. C. R., & Gray, J. A. (2004). Personality Predicts Brain Responses to Cognitive Demands. *The Journal of Neuroscience*, *24*(47), 10636–10641. doi:10.1523/JNEUROSCI.3206-04.2004

Lauriola, M., & Levin, I. P. (2001). Personality traits and risky decision-making in a controlled experimental task: an exploratory study. *Personality and Individual Differences*, *31*(2), 215–226. doi:10.1016/S0191-8869(00)00130-6

Lauriola, M., Panno, A., Levin, I. P., & Lejuez, C. W. (2013). Individual Differences in Risky Decision Making: A Meta-analysis of Sensation Seeking and Impulsivity with the Balloon Analogue Risk Task. *Journal of Behavioral Decision Making*, 20–36. doi:10.1002/bdm.1784

LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cellular and Molecular Neurobiology*, *23*(4-5), 727–738.

Lee, A. K. W., Jerram, M., Fulwiler, C., & Gansler, D. A. (2011). Neural correlates of impulsivity factors in psychiatric patients and healthy volunteers: a voxel-based morphometry study. *Brain Imaging and Behavior*, *5*(1), 52–64. doi:10.1007/s11682-010-9112-1

Lee, V., & Harris, L. (2013). How social cognition can inform social decision making. *Decision Neuroscience*, *7*, 259. doi:10.3389/fnins.2013.00259

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion*, *14*, 473–493.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*(1), 146–159.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. (2015). Emotion and Decision Making. *Annual Review of Psychology*, *66*(1), null. doi:10.1146/annurev-psych-010213-115043

Lerner, J. S., Small, D. A., & Loewenstein, G. (2004). Heart strings and purse strings: Carryover effects of emotions on economic decisions. *Psychological Science*, *15*(5), 337–341. doi:10.1111/j.0956-7976.2004.00679.x

Levy, D. J., & Glimcher, P. W. (2011). Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain. *The Journal of Neuroscience*, *31*(41), 14693–14707. doi:10.1523/JNEUROSCI.2218-11.2011

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027–1038. doi:10.1016/j.conb.2012.06.001

Lighthall, N. R., Mather, M., & Gorlick, M. A. (2009). Acute Stress Increases Sex Differences in Risk Seeking in the Balloon Analogue Risk Task. *PLoS ONE*, *4*(7), e6002. doi:10.1371/journal.pone.0006002

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *The Behavioral and Brain Sciences*, *35*(3), 121–143. doi:10.1017/S0140525X11000446

Loewenstein, G. (1987). Anticipation and the Valuation of Delayed Consumption. *Economic Journal*, *97*(387), 666–84.

Loewenstein, G., & Lerner, J. S. (2003). Loewenstein G, Lerner JS. The role of affect in decision making. In R. J. Davidson, H. Goldsmith, & K. Scherer (Eds.), *Handbook of Affective Science*. Oxford: Oxford University Press.

Logan, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language*. San Diego: Academic Press.

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex in Cognitive Control. *Science*, *288*(5472), 1835–1838. doi:10.1126/science.288.5472.1835

Magid, V., Maclean, M. G., & Colder, C. R. (2007). Differentiating between sensation seeking and impulsivity through their mediated relations with alcohol use and problems. *Addictive Behaviors*, *32*(10), 2046–2061. doi:10.1016/j.addbeh.2007.01.015

Maner, J. K., Richey, J. A., Cromer, K., Mallott, M., Lejuez, C. W., Joiner, T. E., & Schmidt, N. B. (2007). Dispositional anxiety and risk-avoidant decision-making. *Personality and Individual Differences*, *42*(4), 665–675. doi:10.1016/j.paid.2006.08.016

Manning, J., Hedden, T., Wickens, N., Whitfield-Gabrieli, S., Prelec, D., & Gabrieli, J. D. E. (2014). Personality influences temporal discounting preferences: Behavioral and brain evidence. *NeuroImage*, *98*, 42–49. doi:10.1016/j.neuroimage.2014.04.066

Mansouri, F. A., Tanaka, K., & Buckley, M. J. (2009). Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nature Reviews. Neuroscience*, *10*(2), 141–152. doi:10.1038/nrn2538

Martin, S. B., Covell, D. J., Joseph, J. E., Chebrolu, H., Smith, C. D., Kelly, T. H., … Gold, B. T. (2007). Human experience seeking correlates with hippocampus volume:

Convergent evidence from manual tracing and voxel-based morphometry. *Neuropsychologia*, *45*(12), 2874–2881. doi:10.1016/j.neuropsychologia.2007.05.009

Matsuo, K., Nicoletti, M., Nemoto, K., Hatch, J. P., Peluso, M. A. M., Nery, F. G., & Soares, J. C. (2009). A voxel-based morphometry study of frontal gray matter correlates of impulsivity. *Human Brain Mapping*, *30*(4), 1188–1195. doi:10.1002/hbm.20588

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. doi:10.1146/annurev.neuro.24.1.167

Minati, L., Grisoli, M., Franceschetti, S., Epifani, F., Granvillano, A., Medford, N., … Critchley, H. D. (2012). Neural signatures of economic parameters during decision-making: a functional MRI (FMRI), electroencephalography (EEG) and autonomic monitoring study. *Brain Topography*, *25*(1), 73–96. doi:10.1007/s10548-011-0210-1

Minati, L., Grisoli, M., Seth, A. K., & Critchley, H. D. (2012). Decision-making under risk: a graph-based network analysis using functional MRI. *NeuroImage*, *60*(4), 2191–2205. doi:10.1016/j.neuroimage.2012.02.048

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1309–1316. doi:10.1098/rstb.2008.0318

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*(4), 655–663. doi:10.1016/j.neuron.2006.03.040

Miu, A. C., Heilman, R. M., & Houser, D. (2008). Anxiety impairs decision-making: psychophysiological evidence from an Iowa Gambling Task. *Biological Psychology*, *77*(3), 353–358. doi:10.1016/j.biopsycho.2007.11.010

Monterosso, J., Piray, P., & Luo, S. (2012). Neuroeconomics and the study of addiction. *Biological Psychiatry*, *72*(2), 107–112. doi:10.1016/j.biopsych.2012.03.012

Morishima, Y., Schunk, D., Bruhin, A., Ruff, C. C., & Fehr, E. (2012). Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron*, *75*(1), 73–79. doi:10.1016/j.neuron.2012.05.021

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, *1251*, E1–24. doi:10.1111/j.1749-6632.2012.06751.x

Omura, K., Constable, R. T., & Canli, T. (2005). Amygdala gray matter concentration is associated with extraversion and neuroticism. *Neuroreport*, *16*(17), 1905–1908.

O'Shea, J., Taylor, P. C. J., & Rushworth, M. F. S. (2008). Imaging causal interactions during sensorimotor processing. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *44*(5), 598–608. doi:10.1016/j.cortex.2007.08.012

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*(6), 768–774.

Paulus, M. P., & Stein, M. B. (2006). An Insular View of Anxiety. *Biological Psychiatry*, *60*(4), 383–387. doi:10.1016/j.biopsych.2006.03.042

Paulus, M. P., & Yu, A. J. (2012). Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, *16*(9), 476–483. doi:10.1016/j.tics.2012.07.009

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, *13*(4), 160–166. doi:10.1016/j.tics.2009.01.006

Pessoa, L., Padmala, S., Kenzer, A., & Bauer, A. (2012). Interactions between cognition and emotion during response inhibition. *Emotion (Washington, D.C.)*, *12*(1), 192–197. doi:10.1037/a0024109

Phan, K. L., Sripada, C. S., Angstadt, M., & McCabe, K. (2010). Reputation for reciprocity engages the brain reward center. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(29), 13099–13104. doi:10.1073/pnas.1008137107

Phelps, E. A. (2009). The study of emotion in neuroeconomics. In P. W. Glimcher, C. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision Making and the Brain*. New York: Academic Press.

Phelps, E. A., Lempert, K. M., & Sokol-Hessner, P. (2014). Emotion and Decision Making: Multiple Modulatory Neural Circuits. *Annual Review of Neuroscience*, *37*(1), null. doi:10.1146/annurev-neuro-071013-014119

Phillips, M. L., Ladouceur, C. D., & Drevets, W. C. (2008). Neural systems underlying voluntary and automatic emotion regulation: toward a neural model of bipolar disorder. *Molecular Psychiatry*, *13*(9), 829–829. doi:10.1038/mp.2008.82

Pizzagalli, D. A., Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2008). Reduced Hedonic Capacity in Major Depressive Disorder: Evidence from a Probabilistic Reward Task. *Journal of Psychiatric Research*, *43*(1), 76–87. doi:10.1016/j.jpsychires.2008.03.001

Platt, M. L., & Huettel, S. A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nature Neuroscience*, *11*(4), 398–403. doi:10.1038/nn2062

Porcelli, A. J., & Delgado, M. R. (2009). Acute Stress Modulates Risk Taking in Financial Decision Making. *Psychological Science*, *20*(3), 278–283. doi:10.1111/j.1467-9280.2009.02288.x

Porcelli, A. J., Lewis, A. H., & Delgado, M. R. (2012). Acute stress influences neural circuits of reward processing. *Frontiers in Neuroscience*, *6*, 157. doi:10.3389/fnins.2012.00157

Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human Insula Activation Reflects Risk Prediction Errors As Well As Risk. *The Journal of Neuroscience*, *28*(11), 2745–2752. doi:10.1523/JNEUROSCI.4286-07.2008

Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, *17*(1), 15–22. doi:10.1901/jeab.1972.17-15

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556. doi:10.1038/nrn2357

Rick, S., & Loewenstein, G. F. (2008). The Role of Emotion in Economic Behavior. In *Handbook of Emotions* (3rd ed.). Guilford Press. Retrieved from http://dx.doi.org/10.2139/ssrn.954862

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage*, *22*(4), 1694–1703. doi:10.1016/j.neuroimage.2004.04.015

Robinson, O. J., Charney, D. R., Overstreet, C., Vytal, K., & Grillon, C. (2012). The adaptive threat bias in anxiety: Amygdala–dorsomedial prefrontal cortex coupling and aversive amplification. *NeuroImage*, *60*(1), 523–529. doi:10.1016/j.neuroimage.2011.11.096

Robinson, O. J., Letkiewicz, A. M., Overstreet, C., Ernst, M., & Grillon, C. (2011). The effect of induced anxiety on cognition: threat of shock enhances aversive processing in healthy individuals. *Cognitive, Affective & Behavioral Neuroscience*, *11*(2), 217–227. doi:10.3758/s13415-011-0030-5

Robinson, O. J., Overstreet, C., Charney, D. R., Vytal, K., & Grillon, C. (2013). Stress increases aversive prediction error signal in the ventral striatum. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(10), 4129–4133. doi:10.1073/pnas.1213923110

Robinson, O. J., Vytal, K., Cornwell, B. R., & Grillon, C. (2013). The impact of anxiety upon cognition: perspectives from human threat of shock studies. *Frontiers in Human Neuroscience*, *7*, 203. doi:10.3389/fnhum.2013.00203

Rosen, J. B., & Schulkin, J. (1998). From normal fear to pathological anxiety. *Psychological Review*, *105*(2), 325–350. doi:10.1037/0033-295X.105.2.325

Rossi, S., Hallett, M., Rossini, P. M., Pascual-Leone, A., & Safety of TMS Consensus Group. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *120*(12), 2008–2039. doi:10.1016/j.clinph.2009.08.016

Ruchsow, M., Groen, G., Kiefer, M., Hermle, L., Spitzer, M., & Falkenstein, M. (2008). Impulsiveness and ERP components in a Go/Nogo task. *Journal of Neural Transmission*, *115*(6), 909–915. doi:10.1007/s00702-008-0042-7

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews. Neuroscience*, *15*(8), 549–562. doi:10.1038/nrn3776

Ruocco, A. C., Amirthavasagam, S., Choi-Kain, L. W., & McMain, S. F. (2013). Neural correlates of negative emotionality in borderline personality disorder: an activation-likelihood-estimation meta-analysis. *Biological Psychiatry*, *73*(2), 153–160. doi:10.1016/j.biopsych.2012.07.014

Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience*, *7*(5), 499–500. doi:10.1038/nn1223

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science (New York, N.Y.)*, *300*(5626), 1755–1758. doi:10.1126/science.1082976

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, *16*(2), 235–239.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842.

Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, *42*(11), 1435–1446. doi:10.1016/j.neuropsychologia.2004.04.015

Schilling, C., Kühn, S., Romanowski, A., Banaschewski, T., Barbot, A., Barker, G. J., … Consortium, the I. (2013). Common structural correlates of trait impulsiveness and perceptual reasoning in adolescence. *Human Brain Mapping*, *34*(2), 374–383. doi:10.1002/hbm.21446

Schmitz, A., & Grillon, C. (2012). Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test). *Nature Protocols*, *7*(3), 527–532. doi:10.1038/nprot.2012.001

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality & Social Psychology Bulletin*, *34*(8), 1096–1109. doi:10.1177/0146167208317771

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*, *275*(5306), 1593–1599.

Schutter, D. J. L. G., Van Honk, J., Koppeschaar, H., & Kahn, R. (2002). Cortisol and reduced interhemispheric coupling between the left prefrontal and the right parietal cortex. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *14*(1), 89–90.

Schwabe, L., Tegenthoff, M., Höffken, O., & Wolf, O. T. (2012). Simultaneous Glucocorticoid and Noradrenergic Activity Disrupts the Neural Basis of Goal-Directed Action in the Human Brain. *The Journal of Neuroscience*, *32*(30), 10146–10155. doi:10.1523/JNEUROSCI.1304-12.2012

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*(3), 513–523. doi:10.1037/0022-3514.45.3.513

Sescousse, G., Caldú, X., Segura, B., & Dreher, J.-C. (2013). Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *37*(4), 681–696. doi:10.1016/j.neubiorev.2013.02.002

Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews. Neuroscience*, *12*(3), 154–167. doi:10.1038/nrn2994

Sharp, C., Monterosso, J., & Montague, P. R. (2012). Neuroeconomics: a bridge for translational research. *Biological Psychiatry*, *72*(2), 87–92. doi:10.1016/j.biopsych.2012.02.029

Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social Cognitive and Affective Neuroscience*, *8*(3), 341–350. doi:10.1093/scan/nss002

Sokol-Hessner, P., Hartley, C. A., Hamilton, J. R., & Phelps, E. A. (2014). Interoceptive ability predicts aversion to losses. *Cognition & Emotion*, 1–7. doi:10.1080/02699931.2014.925426

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences*, *106*(13), 5035–5040. doi:10.1073/pnas.0806761106

Spampinato, M. D., Wood, P. D., De Simone, M. D., & Grafman, P. D. (2009). Neural Correlates of Anxiety in Healthy Volunteers: A Voxel-Based Morphometry Study. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *21*(2), 199–205.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Stanton, S. J., Reeck, C., Huettel, S. A., & LaBar, K. S. (2014). Effects of induced moods on economic choices. *Judgment and Decision Making*, *9*(2), 167–175.

Stein, M. B., Jang, K. L., & Livesley, W. J. (1999). Heritability of anxiety sensitivity: a twin study. *The American Journal of Psychiatry*, *156*(2), 246–251.

Stoel, R. D., De Geus, E. J. C., & Boomsma, D. I. (2006). Genetic analysis of sensation seeking with an extended twin design. *Behavior Genetics*, *36*(2), 229–237. doi:10.1007/s10519-005-9028-5

Stokes, M. G., Chambers, C. D., Gould, I. C., Henderson, T. R., Janko, N. E., Allen, N. B., & Mattingley, J. B. (2005). Simple Metric For Scaling Motor Threshold Based on Scalp-Cortex Distance: Application to Studies Using Transcranial Magnetic Stimulation. *Journal of Neurophysiology*, *94*(6), 4520–4527. doi:10.1152/jn.00067.2005

Sweitzer, M. M., Allen, P. A., & Kaut, K. P. (2008). Relation of individual differences in impulsivity to nonclinical emotional decision making. *Journal of the International Neuropsychological Society*, *14*(05), 878–882. doi:10.1017/S1355617708080934

Tackett, J. L., Waldman, I. D., Van Hulle, C. A., & Lahey, B. B. (2011). Shared genetic influences on negative emotionality and major depression/conduct disorder

comorbidity. *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*(8), 818–827. doi:10.1016/j.jaac.2011.05.007

Talmi, D., Hurlemann, R., Patin, A., & Dolan, R. J. (2010). Framing effect following bilateral amygdala lesion. *Neuropsychologia*, *48*(6), 1823–1827. doi:10.1016/j.neuropsychologia.2010.03.005

Tan, H. B., & Forgas, J. P. (2010). When happiness makes us selfish, but sadness makes us fair: Affective influences on interpersonal strategies in the dictator game. *Journal of Experimental Social Psychology*, *46*(3), 571–576. doi:10.1016/j.jesp.2010.01.007

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: the effects of specific emotions on information processing. *Journal of Personality and Social Psychology*, *81*(6), 973–988.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science*, *315*(5811), 515–518. doi:10.1126/science.1134239

Tracey, I., & Mantyh, P. W. (2007). The Cerebral Signature for Pain Perception and Its Modulation. *Neuron*, *55*(3), 377–391. doi:10.1016/j.neuron.2007.07.012

Treadway, M. T., & Zald, D. H. (2011). Reconsidering anhedonia in depression: lessons from translational neuroscience. *Neuroscience and Biobehavioral Reviews*, *35*(3), 537–555. doi:10.1016/j.neubiorev.2010.06.006

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. doi:10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1986). Rational Choice and the Framing of Decisions. *The Journal of Business*, *59*(4), S251–S278. doi:10.2307/2352759

Van Gestel, S., & Van Broeckhoven, C. (2003). Genetics of personality: are we making progress? *Molecular Psychiatry*, *8*(10), 840–852. doi:10.1038/sj.mp.4001367

Von Neumann, J., & Morgenstern, O. (2007). *Theory of Games and Economic Behavior (Commemorative Edition)*. Princeton University Press.

Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A., & Ochsner, K. N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron*, *59*(6), 1037–1050. doi:10.1016/j.neuron.2008.09.006

Walther, H. (2003). Normal-randomness expected utility, time preference and emotional distortions. *Journal of Economic Behavior & Organization*, *52*(2), 253–266. doi:10.1016/S0167-2681(03)00005-2

Xu, P., Gu, R., Broster, L. S., Wu, R., Dam, N. T. V., Jiang, Y., … Luo, Y. (2013). Neural Basis of Emotional Decision Making in Trait Anxiety. *The Journal of Neuroscience*, *33*(47), 18641–18653. doi:10.1523/JNEUROSCI.1253-13.2013

Yamasue, H., Abe, O., Suga, M., Yamada, H., Inoue, H., Tochigi, M., … Kasai, K. (2008). Gender-Common and -Specific Neuroanatomical Basis of Human Anxiety-Related Personality Traits. *Cerebral Cortex*, *18*(1), 46–52. doi:10.1093/cercor/bhm030

Yarkoni, T. (2015). Neurobiological substrates of personality: A critical overview. In M. Mikulincer, P. R. Shaver, M. L. Cooper, & R. J. Larsen (Eds.), *APA handbook of personality and social psychology, Volume 4: Personality processes and individual differences* (pp. 61–83). Washington, DC, US: American Psychological Association.

Yeung, C. W. M., & Wyer, Jr., R. S. (2004). Affect, Appraisal, and Consumer Judgment. *Journal of Consumer Research*, *31*(2), 412–424. doi:10.1086/422119

Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences*, 200914826. doi:10.1073/pnas.0914826107

Zermatten, A., Van der Linden, M., d' Acremont, M., Jermann, F., & Bechara, A. (2005). Impulsivity and decision making. *The Journal of Nervous and Mental Disease*, *193*(10), 647–650.

Zuckerman, M. (1994). *Behavioral Expressions and Biosocial Bases of Sensation Seeking*. Cambridge, United Kingdom: Cambridge University Press.

Zuckerman, M. (1996). Item revisions in the Sensation Seeking Scale Form V (SSS-V). *Personality and Individual Differences*, *20*(4), 515. doi:10.1016/0191-8869(95)00195-6

Appendix

A Framing effects on risky decision-making: Investigating a causal role for the dorsolateral prefrontal cortex with TMS

Friederike Meyer[1], Daniela M. Pfabigan[2], Christian C. Ruff[1]

[1]Laboratory for Social and Neural Systems Research (SNS-Lab), Department of Economics, University of Zurich, Zurich, Switzerland

[2]Social, Cognitive and Affective Neuroscience Unit, Faculty of Psychology, University of Vienna, Vienna, Austria

A.1 Abstract

Contrary to assumptions of standard economic theories, human decision-making is commonly influenced by contextual factors and systematic cognitive biases. Description invariance, a hallmark of normative decision-making theory, can be violated, for instance, by the framing effect: Different, but logically equivalent descriptions of the same choice set lead to differing preferences. Whereas a positively framed risky prospect elicits risk aversion, a negatively framed prospect of the same choice set is associated with risk-seeking, leading to enhanced reflection effects.

Since previous neuroimaging studies identified the dorsolateral prefrontal cortex (DLPFC) as a neural correlate of the inhibition of the framing effect, we sought to probe for the causal role of DLPFC in susceptibility to framing by means of repetitive transcranial magnetic stimulation (rTMS). Although the previous studies suggested a substantial role of DLPFC in framing, the current study could not confirm its functional relevance for inhibition of financial risky choice framing effects using unilateral disruption of either left or right DLPFC. It might be that the causal role of DLPFC in inhibition of framing susceptibility has been overestimated due to results of previous correlative neuroimaging studies. In order to yield firm conclusions on the functional relevance of DLPFC for framing effects, methodological modifications and experimental approaches for future studies are discussed.

A.2 Introduction

Changes in risk preferences due to cognitive biases are a common observation in psychological research and behavioral economics (Tversky & Kahneman, 1974).

The dependence of risk preferences on the description of a risky choice problem is called the framing effect (Tversky & Kahneman, 1986) and represents a major violation of description invariance, a major axiom of rational choice theory: According to rational choice theory equivalent descriptions of the same choice problem should lead to identical preferences. However, humans actually tend to behave risk-averse when risky prospects are presented in a positive frame, whereas increased risk-seeking is observed when the same options are embedded in a negative context. For instance, financial decisions are

substantially influenced by their description as a loss or gain relative to a reference point, which often is the status quo (Tversky & Kahneman, 1986). Prospect theory (Kahneman & Tversky, 1979) explains the framing effect by the specific properties of the value function, which is concave for gains and convex for losses (implying diminishing sensitivity for both value domains), and the reference point established by the frame. Due to the form of the value function, decision-makers oftener opt for a sure option evaluated as a gain rather than choosing a risky option of equal expected value. Conversely, they tend to prefer the risky option over a sure option of equal expected value, which is perceived as a loss.

From a psychological perspective, emotions are thought to play a key role in the framing effect. In particular, dual process theories (De Neys, 2006; Evans, 2003; Sloman, 1996) offer an explanation for this decision-making bias by proposing the dominance of an intuitive/affective system (System 1) over a cognitive/executive system (System 2) (Cassotti et al., 2012; Kahneman & Frederick, 2007). The dominance of System 1, which is characterized by rapid and automatic operations, leads to the employment of affective heuristics causing susceptibility to framing. The computationally demanding and slow operations of System 2, however, are assumed to be able to override the affective heuristic-based responses generated by System 1. Thus, according to dual process theories, susceptibility to framing is hypothesized to be caused by the dominance of System 1, whereas overcoming the behavioral tendencies imposed by frames is characterized by the dominance of System 2 over System 1.

This "affective heuristics hypothesis" of framing has been supported by recent neuroimaging studies. While increased amygdala activation was observed for frame-congruent decisions (choices made *according to* the behavioral tendency imposed by the frame), frame-incongruent decisions (choices made *against* the behavioral heuristics imposed by the frame) were correlated with increased activation in DLPFC and anterior cingulate cortex (ACC) (De Martino, Kumaran, Seymour, & Dolan, 2006). This activation pattern was interpreted as empirical evidence for an emotion-driven framing effect, which can be overcome by the implementation of cognitive control of these behavioral tendencies via DLPFC and ACC (De Martino et al., 2006; Gonzalez, Dana, Koshino, & Just, 2005). This interpretation is clearly in line with the well-established role of these brain regions in the implementation and maintenance of top-down cognitive control over behavior and decision-making (Coutlee & Huettel, 2012; MacDonald, Cohen, Stenger, & Carter, 2000; Miller & Cohen, 2001). Both right and left DLPFC have been observed to be involved in computations relevant to those functions (Fecteau et al., 2007; Figner et al., 2010; Hare,

Camerer, & Rangel, 2009; Knoch et al., 2006). However, direct causal evidence for the role of dorsolateral prefrontal brain structures in resistance to framing is missing until date, since the aforementioned neuroimaging studies offer solely correlational evidence for the involvement of the DLPFC in resistance to framing. While functional magnetic resonance imaging (fMRI) identifies correlational links between brain activity and behavior, transcranial magnetic stimulation (TMS) allows establishing a causal link between brain function and behavior. Using fMRI, we cannot conclude whether a brain region is pivotal in the generation of a specific behavior or merely coactivated by a neural network causally involved in the given behavior. Specifically, DLPFC may implement computations that are correlated with resistance to frames, but that do not contribute to framing resistance itself. A hypothesis about a brain region's causal relevance for a behavioral manifestation – like the framing effect – provided by correlative tools such as fMRI can be complemented by TMS.

Given the involvement of the DLPFC in frame-incongruent decisions found by De Martino and colleagues (2006), we hypothesized that temporary disruption of DLPFC functioning will lead to changes in susceptibility to framing. Specifically, taking into account the involvement of bilateral DLPFC in frame-incongruent choices reported by De Martino et al. (2006), we tested whether temporary perturbation of DLPFC functioning would cause increased frame-congruent decisions due to the reduction of cognitive control over this automatic behavioral tendency. To this end, we employed a TMS-compatible risky framing task similar to that used by De Martino et al. (2006) after participants had undergone continuous theta-burst stimulation (cTBS; Huang et al., 2009; Huang, Edwards, Rounis, Bhatia, & Rothwell, 2005). cTBS is thought to interrupt cortical activity in targeted brain regions for up to 20 minutes (Huang et al., 2009, 2005) after stimulation, and was applied to either the left or right DLPFC (treatment conditions) or to subject's vertex (control condition). Importantly, subjects were randomly assigned to the three experimental conditions. Because of the well-known individual differences in susceptibility to the framing effect (Levin, Gaeth, Schreiber, & Lauriola, 2002), each participant completed the same risky decision-making task twice, once with and once without cTBS, in order to get within-subject measurements of the TMS effect on framing susceptibility relative to a purely behavioral baseline measurement.

A.3 Methods

Participants. 67 (36 females; 18 - 35 years, mean age [std] = 23.08 years [3.54]) right-handed adults participated in the experiment. In total, 4 of these participants had to be excluded due to technical issues (n = 2) during the experiment or previous experience with the behavioral task (n = 2). All subjects gave written informed consent to the experimental procedure. Their eligibility for participation in a TMS study was carefully assessed, that is, subjects with a history of neurologic or psychiatric diseases (assessed via a pre-experimental questionnaire) were excluded from the study. Subsequently, subjects were randomly and evenly assigned to the three experimental conditions, resulting in a final group size of n = 21 for left and right DLPFC as well as the vertex group. The study was approved by the ethics committee of the Canton Zürich.

Experimental Timeline. The study consisted of two sessions scheduled for separate days. The first appointment served to conduct the TMS experiment, which aimed at investigating the neurobiological substrates of the framing effect. The second appointment, conducted approx. 14 days later, entailed a purely behavioral session without TMS treatment. During this second session, the participants completed the same behavioral financial risky choice framing task as in the TMS experiment, in order to derive a baseline estimate of the individual framing effect and to enable within-subject comparisons. In order to reduce consistency or memory effects, the paradigm was modified for the second session (see section *Behavioral task*).

Instruction and Questionnaire phase. Upon arrival at the laboratory, subjects were informed about the experimental procedure and read detailed information about the experimental task. Five practice trials ensured that the participants understood the behavioral task and were familiar with the visual displays (see section "*Behavioral task"*). Additionally, participants were asked to complete a series of psychological questionnaires, which covered a broad range of personality dimensions. The purpose of this questionnaire battery was to investigate the potential influence of personality characteristics on susceptibility to framing. To this end, participants completed the Barratt Impulsiveness Scale (BIS-11; Patton, Stanford, & Barratt, 1995), the State-Trait-Anxiety-Inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983), the NEO-FFI (Borkenau & Ostendorf, 1993) for assessment of personality characteristics such as neuroticism, extraversion and openness, and the Need for Cognition Scale (Bless, Wänke, Bohner,

Fellhauer, & Schwarz, 1994; Cacioppo & Petty, 1982), which measures dispositional differences in the intensity of cognitive processing and has been found to correlate with the interindividually varying degree of susceptibility to framing manipulations (Smith & Levin, 1996).

TMS procedure. Before the experiment, anatomical magnetic resonance brain images were obtained for each subject. These high-resolution T1-weighted 3D-TFE images were acquired using a 3T whole-body MR Scanner (Philips Medical Systems, Best, The Netherlands) at the SNS laboratory of the University of Zürich (TR: 7500 ms; TE = 3.5 ms; FA = 8°; FOV 250 × 250 mm; voxel size 1.04 × 1.04 × 0.6 mm; 301 sagittal slices).

Sites for TMS application were based on the study by De Martino et al. (2006). They observed increased DLPFC activation when subjects resisted the behavioral tendency imposed by a gain or loss frame. That is, enhanced bilateral DLPFC activation was observed when subjects chose the lottery in the gain frame and the sure option in loss frame ($[S_{LF} + L_{GF}] > [S_{GF} + L_{LF}]$). In the present experiment, TMS was either applied to participants' left or right DLPFC coordinates from this study (left PFC: x = -48, y = 18, z = 24; right PFC: x = 56, y = 18, z = 28 [MNI space], Figure A-1) or over the subject's vertex, which was defined as the conjunction point of the left and right central fissures in the interhemispheric fissure of each subject's individual brain.

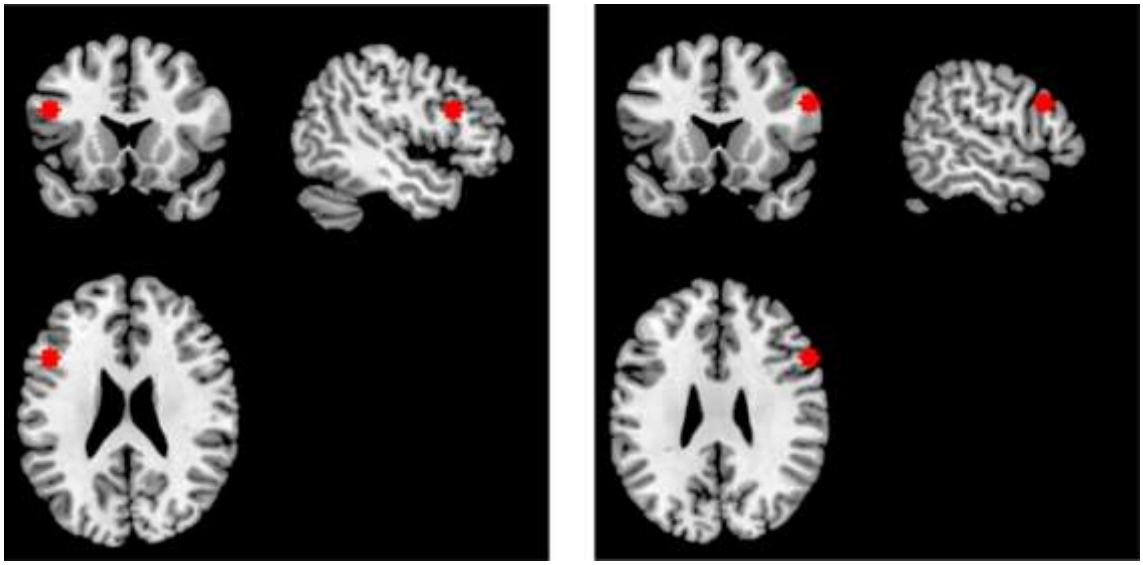

Figure A-1: TMS sites. cTBS was either applied to participants' left or right DLPFC coordinates (left DPFC: x = -48, y = 18, z = 24; right DPFC: x = 56, y = 18, z = 28 [MNI space]), left and right panel, respectively. In the control condition, stimulation was applied over the subject's vertex, which was defined as the conjunction point of the left and right central fissures in the interhemispheric fissure of each subject's individual brain (not shown/graphically displayed here).

To determine the left and right stimulation site for each subject, his anatomical scan was normalized into MNI space using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/). The individual target coordinates for TMS were determined in the normalized scan and backprojected from MNI into subject space applying inverse normalization. These individual coordinates were then marked on the participant's structural image using Brainsight (Rogue Research Inc., Montreal, Canada) frameless stereotactic neuronavigation. The vertex was determined based on each participant's individual neuroanatomy (see exact procedure above) and marked on each subject's T1 scan.

In order to determine subjects' individual active motor threshold (AMT), stimulation of the left M1 was carried out over the optimal location for stimulation of the right first dorsal interosseus (FDI) muscle. We used a 70 mm-figure of eight-coil placed tangentially over the participant's right primary motor cortex with 45° backwards and laterally over the hot-spot of the left hand FDI in order to stimulate M1. To obtain active AMTs, used in cTBS for safety reasons, muscle force was generated by participants squeezing forefinger and thumb to activate their FDI muscle at approximately 25 % of maximum force. The AMT was thus measured during 25% of the maximum voluntary contraction of the FDI muscle, and was defined as the lowest stimulator output that elicited a visible motor twitch in the FDI muscle and motor evoked potentials of at least 200 µV in at least in 5 out of 10 trials. In particular, a Magstim Rapid^2 stimulator (The Magstim Company Ltd, Whitland, U.K.) was used to stimulate cortical regions. MEPs were registered using surface electrodes placed over participants' left FDI. cTBS was applied at an intensity of 80% AMT using approved protocols (Huang et al., 2005; Katayama & Rothwell, 2007). We delivered cTBS with the same Magstim Rapid^2 stimulator (The Magstim Company Ltd, Whitland, U.K.) and a 70 mm figure-of-eight coil also used for AMT measurements. The cTBS protocol consisted of 40 s continuous triplets of 50 Hz that repeat every 200 ms. In total, 600 pulses were delivered. Coil control was performed with Brainsight, based on the individually prepared T1-weighted structural scans. For left and right prefrontal cortex (PFC) stimulation, the coil was positioned with the handle pointing in an upward vertical orientation, at a tangential position to the participant's head. For the control condition, the coil was placed over the vertex with the handle pointing in a posterior direction parallel to the midline. Throughout all three conditions, the appropriate coil position in relation to the subject's head was constantly monitored by the neuronavigation system.

TMS parameters. The group-averaged AMT (± standard error of the mean [SEM]) were 51.4 ± 2.1 %, 44.7 ± 1.5 and 45.4 ± 1.3 and 44 ± 6.8 % of the maximum stimulator output (MSO) in the vertex, left group and right TMS group, respectively. cTBS intensities (± SEM) were therefore 40 ± 1.4 % and 36 ± 1.3 %, 36.1 ± 1.1 of the MSO in vertex, left and right TMS group, respectively.

Behavioral paradigm. The framing task was conducted in Matlab R2012a (The MathWorks Inc, Natick, Massachusetts, USA) using Cogent 2000 (http://www.vislab.ucl.ac.uk). All experimental stimuli were presented on a computer monitor against a black background.

Each trial started with an endowment phase, during which an endowment (E) for the respective trial was presented on the screen for 1.5 s. This endowment transferred a specific number of points to the participant. Subjects were instructed that these points would be used for the subsequent choices and that the points gained in four choices randomly drawn after the experiment would be converted into CHF at a fixed rate (15 points = 1 CHF in the TMS session, 40 points = 1 CHF in the behavioral session). These payments were real so that the participants' choices were fully incentive compatible.

The endowment phase was followed by the decision stage lasting a maximum of 4 s. In this phase, the subjects were asked to decide between two available options. They had to choose either to play a lottery (L) with varying probabilities of winning the whole endowment [P(WIN)] or to take a fraction of the endowment for sure (S). For each target trial the expected values of the L and S options were equivalent, that is, E*P(WIN) = S. Importantly, S was either framed as a gain ("KEEP", gain frame [GF]) or a loss ("LOSE", loss frame [LF]). The probabilities of winning the whole endowment amount were 0.2, 0.4, 0.6, and 0.8. Each of these values was combined with 14 different endowment levels: 25, 35, 45, 50, 60, 70, 75, 85, 95, 100, 110, 120, 125, and 130 points. Each of these combinations was shown in the GF as well as LF, resulting in a total of 112 target trials (comprising 56 GF and 56 LF trials). To monitor subjects' attention to the task and to divert attention from the equivalence of expected outcomes of L and S in target trials, we also included two types of catch trials. These mainly differed from the target trials by a marked inequivalence of expected outcome of L and S (E*P(WIN) ≠ S). Whereas no mathematically preferable option existed for the target trials, a non-ambiguously preferable option was given in the catch trials. In catch trials of type 1 (Catch_1), the probabilities of winning the whole endowment were 0.05 and 0.95. In this kind of trial type, the sure amount was calculated as S = E*0.5. The set-up of Catch_1 trials resulted in lottery-

weighted or sure option-weighted trials for trials with a probability of winning the whole endowment of 0.05 or 0.95, respectively. The probabilities of catch trials of type 2 (Catch_2) were 0.4 and 0.6 and the sure amount was always calculated as S = E*0.9. Thus, in this kind of trial, the sure option was mathematically always preferable. Seven different endowment levels were used for the catch trials: 20, 30, 40, 90, 105, 115, and 135 points. Equivalent to the target trials, each [E, P(WIN)] catch trial combination was shown in both frames, leading to a total of 56 catch trials (28 Catch_1 trials, 28 Catch_2 trials). Target and catch trials were presented in a pseudorandom order, with the restriction that maximally 4 trials of the same trial type were sequentially presented. Subjects were allowed a maximum reaction time of 4 s. Immediately after choosing the preferred option via the left or right arrow key (preferred option on right screen half: right arrow press requested, preferred option on left screen half: left arrow press requested), a fixation cross was presented for 2 s (ITI) (see Figure A-2 for trial setup).

At the beginning of the TMS session, the spatial locations of S and L option (left or right half of the screen) were pseudorandomly determined across participants and kept constant across the whole experimental session. Notably, during the second session, the two options were presented in reverse order, while other aspects of the experiment were kept constant.

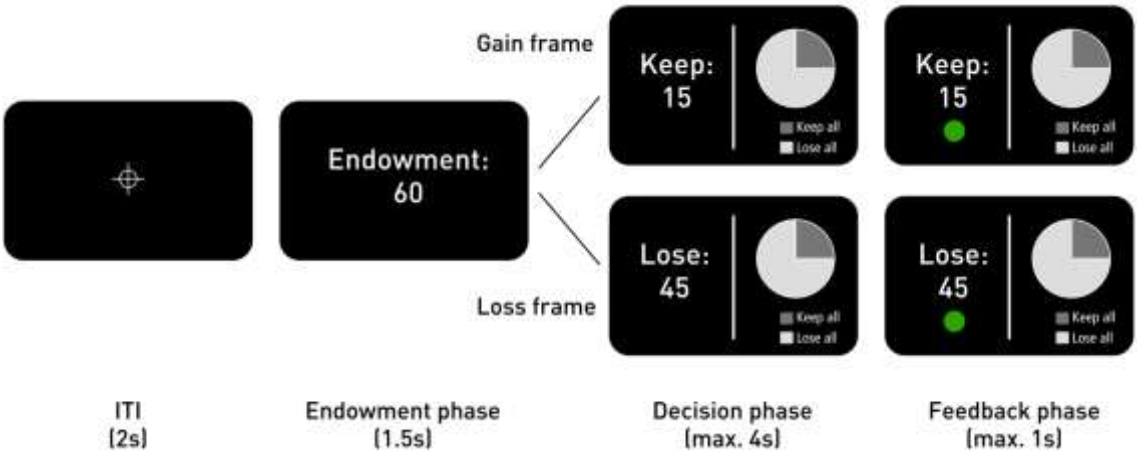

Figure A-2: Trial sequence. Each trial started with a central fixation cross displayed, followed by the endowment stage. During this stage, varying amounts of endowments were presented to the participants during this stage. Next, the decision screen was presented (exact display duration dependent on the participants' response latencies). Subjects had to decide between a lottery (L) and sure (S) option. They could either take a fraction of the initial endowment for sure (S) or play a lottery with varying probabilities of winning the whole initial endowment (L). L was represented by a pie chart depicting the probabilities of losing and winning the whole endowment. Importantly, S was framed either as a gain ("KEEP in the gain frame) or as a loss ("LOSE in the loss frame). Immediately after subjects had chosen their preferred option, their choice was visually confirmed by means of a green circle, which was displayed below the chosen option.

Payment routine. To incentivize behavior and promote subjects' deliberate decision-making, the final payout was determined based on each subject's decisions made during the two experimental sessions. That is, for each session, four decisions were randomly chosen and implemented for real. The resulting sum was added to the fixed participation fee (100 CHF) and paid out at the end of the second session. In order to avoid feedback effects, subjects were informed about their total earnings in both sessions after completion of the second session.

Behavioral data analysis. Data analysis was performed with Matlab R2012a (The MathWorks Inc, Natick, Massachusetts, USA) and STATA 12 (StataCorp LP, College Station, Texas, USA).

Consistent with the literature, the framing effect (FE) was calculated as the difference in lottery choices between LF and GF ($FE = L_{LF} - L_{GF}$). Responses that were congruent with the behavioral tendency imposed by the frame (L_{LF}, S_{GF}) were labeled as frame-congruent, whereas responses indicating resistance to the frame (S_{LF}, L_{GF}) were named frame-incongruent.

In order to investigate the effects of TMS on decision-making, two different types of regression analysis were performed. First, we conducted linear regression analysis of gambling frequencies (i.e., proportion of trials in which L was chosen) to investigate the effects of different experimental factors (e.g., TMS group, frame type (LF, GF), probability of winning level) on subjects' choices. Second, in order to perform an analysis with increased sensitivity to single responses, we performed multiple logistic regression analysis of single-trial choice data. The regression model always contained TMS group (left PFC, right PFC, and vertex) and frame type as independent variables. Different logistic regressions modeled the influence of probability of winning level, endowment amount and personality characteristics as well as their interactions on choices.

A.4 Results

Catch trials

As a very first step, we examined whether subjects exerted continuous engagement with the task. For this purpose, we analyzed Catch_1 and Catch_2 trials. This analysis confirmed that subjects spent continuous attention to the task in both the TMS and Post-TMS sessions. In Catch_1 trials, participants in all treatment conditions showed higher gambling rates in

lottery- weighted than in sure option-weighted trials (see Figure A-3 for mean gambling rates). These results were confirmed by a logistic regression of probability of lottery versus sure option response, which revealed significantly increased probability of gambling in lottery-weighted trials across all three groups (Table A-1). Groups were equally accurate in their choices, as no significant TMS group × weight (lottery-/sure option-weighted) interaction was found.

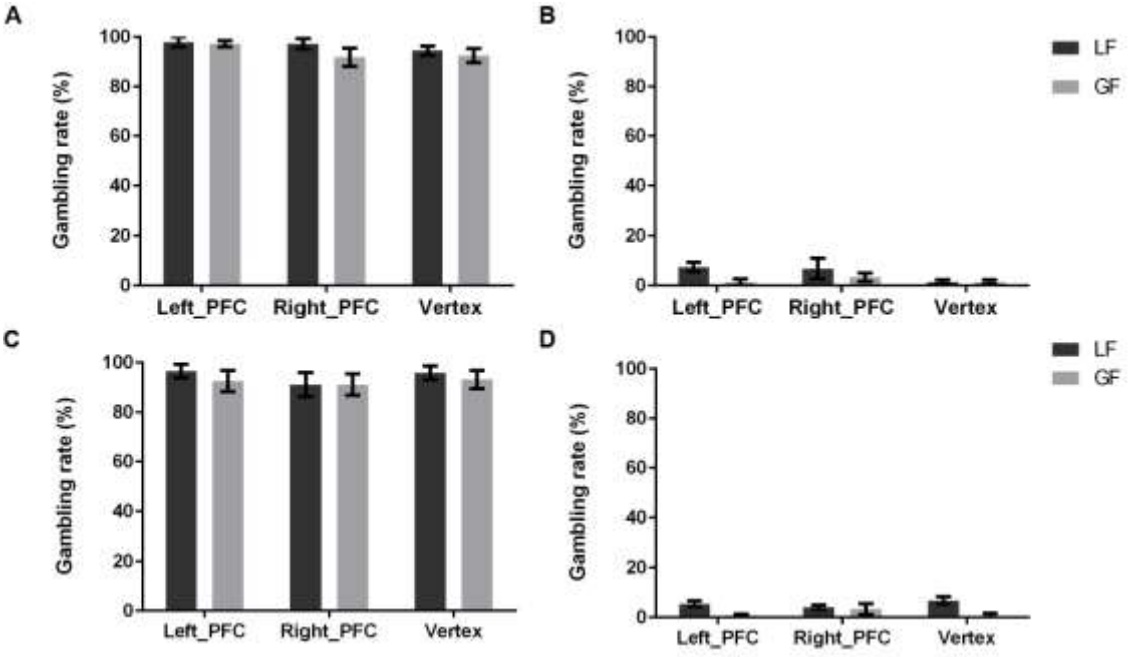

Figure A-3: Analysis of Catch_1 trials: Subjects spent continued attention to the task. A: Shown are gambling rates for lottery-weighted Catch_1 trials in the TMS session. The high-gambling rates for lottery-weighted Catch_1 trials confirmed subjects' continued attention to the task during the TMS session, as they mostly chose the lottery option, which had the higher expected outcome in this kind of trial. B: Shown are gambling rates for sure option-weighted Catch_1 trials, i.e., trials in which the sure option had the higher expected value, in the TMS session. Subjects indeed preferred the sure option in Catch_1, which led to low gambling rates for Catch_1 trials in the TMS session. C: Gambling rates for lottery-weighted Catch_1 trials in the Post-TMS session. High-gambling rates for lottery-weighted Catch_1 trials were also found in the Post-TMS session. D: Shown are gambling rates for sure option-weighted Catch_1 trials in the Post-TMS session. Low gambling rates for Catch_1 trials were also observed in the Post-TMS session. These results evidenced participants' attention to the task during the Post-TMS session. Error bars: SEM

Table A-1: Continuous attention spent to the task reflected by choice behavior in Catch_1 trials

	Variables	**TMS session**	**Post-TMS session**
Response	Left PFC	1.203	-0.438
		(0.653)	(0.675)
	Right PFC	1.357*	-0.252
		(0.73)	(0.827)
	Weight (L, S)	6.949***	5.851***
		(0.653)	(0.777)
	Left PFC × weight (L, S)	-0.166	0.42
		(0.95)	(1.138)
	Right PFC × weight (L, S)	-1.174	-0.27
		(1.016)	(1.165)
	Constant	-4.277***	-2.996***
		(0.584)	(0.505)

Note: Logistic regression of Catch_1 trial responses Subjects spent continuous attention to the task since significantly higher probability of gambling was found in lottery (L)- compared to sure option (S)-weighted trials. Subjects were highly accurate in noticing the imbalanced expected outcomes of L versus S in Catch_1 trials in both TMS and Post-TMS session. Shown are regression coefficients and robust standard errors (SE; in parentheses) clustered by subject. Note: coding scheme: S = 0, L = 1. (Vertex group: baseline). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The Catch_2 trials had the main purpose of distracting participants from the equivalence of expected outcomes of the target trials. The analysis of these sure option-weighted trials confirmed the participants' attention to the task by revealing low gambling rates in both LF and GF across all TMS groups (for mean gambling rates see Figure A-4). Nevertheless, a logistic regression analysis on gambling responses in Catch_2 trials revealed a significant framing effect in the TMS session. That is, despite the superiority of the sure option in those trials, subjects' were still susceptible to the behavioral tendency imposed by the frame. However, in the Post-TMS session, no significant framing effect was found (Table A-2).

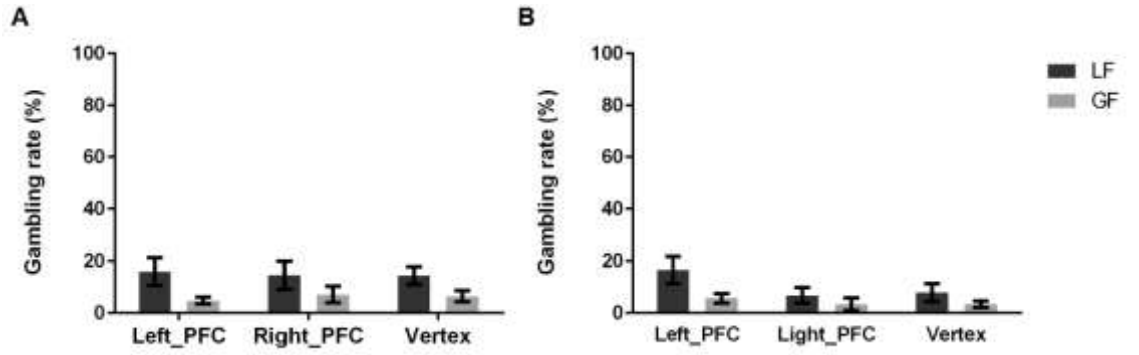

Figure A-4: Gambling rates for sure option-weighted Catch_2 trials in TMS session (A) and Post-TMS session (B). Subjects were able to notice the dominance of the sure option in Catch_2 trials, reflected in low gambling rates for those trials. Error bars: SEM

Table A-2: Continuous attention spent to the task reflected in low gambling rates in Catch_2 trials

	Variables	TMS session	Post-TMS session
Response	Left PFC	-0.323	0.452
		(0.453)	(0.488)
	Right PFC	0.107	0.001
		(0.588)	(0.825)
	Frame	0.885***	0.883
		(0.311)	(0.586)
	Frame × Left PFC	0.452	-0.521
		(0.574)	(0.695)
	Frame × Right PFC	-0.107	-0.154
		(0.549)	(0.719)
	Constant	-2.672***	-3.346***
		(0.349)	(0.38)

Note: Logistic regression of choices in Catch_2 trials. Catch_2 trials had the main purpose of distracting subjects from the expected outcome equivalence of the target trials. In these trials, S (S = 0.9*E) was always preferable to L (L = 0.4 and 0.6*E). Although subjects mostly preferred the lottery, the effect of framing on decision-making, was still observable at least in the TMS session. In particular, whereas a significant framing effect was found in the TMS session, no significantly increased probability to choose L in the LF compared to GF was found in the Post-TMS session. Shown are regression coefficients and robust SE (in parentheses) clustered by subject. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Target trials

To investigate the modulation of the framing effect by cTBS of right or left DLPFC, we tested for differences in the susceptibility to framing during the target trials (characterized by equivalent expected value for the S and L options) between the three TMS groups in the

TMS and the Post-TMS session. Visual inspection of mean gambling rates in TMS and Post-TMS session did not reveal prominent differences between TMS groups (Figure A-5).

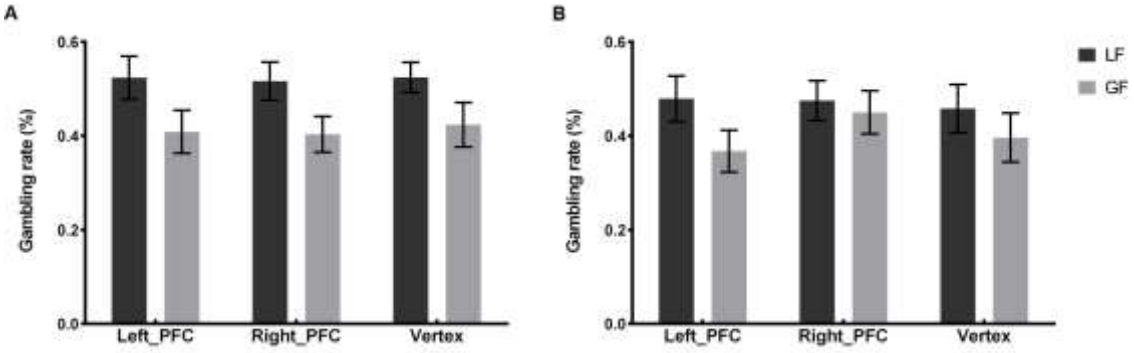

Figure A-5: Mean target trial gambling rates in LF and GF for each TMS group during the TMS and Post TMS session. A: Mean gambling rates (SEM) in TMS session: Left PFC: LF = 0.53 (0.05), GF = 0.4 (0.05); Right PFC: LF = 0.52 (0.04), GF = 0.41 (0.04); Vertex: LF = 0.52 (0.03), GF = 0.42 (0.05); B: Mean gambling rates in LF and GF for each TMS group in the Post-TMS session: Left PFC: LF = 0.48 (0.05), GF = 0.37 (0.04); Right PFC: LF = 0.48 (0.04), GF = 0.45 (0.05); Vertex: LF = 0.45 (0.05), GF = 0.4 (0.05)

This observation was statistically verified by linear regression of participants' gambling rates on frame type and TMS group in TMS and Post-TMS session. This analysis revealed a significant framing effect in the TMS session, i.e., significantly higher gambling rates in LF relative to GF. However, we did not observe a significant effect of TMS group on gambling behavior, nor did this analysis reveal a significant interaction effect between frame type and TMS group. The linear regression analysis of gambling rates in the Post-TMS session, which was supposed to serve as a baseline measurement of framing effects, failed to show susceptibility to framing. Gambling rates in LF were numerically increased compared to the GF during the Post-TMS session, but these effects failed to reach statistical significance (Table A-3). This may possibly reflect learning or memory effects from the prior TMS session.

Table A-3: No effect of cTBS of right or left DLPFC on framing susceptibility

Variables	TMS session	Post-TMS session
Left PFC	-0.0148	-0.0288
	(0.0662)	(0.0673)
Right PFC	-0.02	0.0536
	(0.061)	(0.0673)
Frame	0.101***	0.0616
	(0.0378)	(0.0673)
Frame × Left PFC	0.0142	0.0498
	(0.0603)	(0.0952)
Frame × Right PFC	0.0122	-0.0366
	(0.0566)	(0.0952)
Constant	0.424***	0.396***
	(0.0474)	(0.0476)

Note: Linear regression of gambling rates in target trials. Subjects showed significantly increased gambling in LF relative to GF, reflecting successful induction of risky choice framing effect. No significant TMS group differences or interaction between frame type and TMS group were found. Coding scheme: Frame type: GF = 0, LF = 1; vertex group served as baseline. Shown are regression coefficients and robust SE (in parentheses) clustered by subject. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In order to examine participants' gambling behavior with a measure more sensitive to individual responses, logistic regression of responses (lottery or safe option chosen) were performed for the TMS and Post-TMS session. This model comprised the subjects' responses (L and S) as dependent variable as well as TMS group (left PFC, right PFC, and vertex), and frame type (LF, GF) as independent variables as well as their interactions. For the TMS session, the analysis revealed a significant effect of frame type on choice. That is, the probability of choosing the lottery over the safe option was increased in LF relative to GF. We observed neither an influence of TMS group on responses nor significant interactions between TMS group and frame type.

The same logistic regression was conducted on the Post-TMS session data and confirmed the result of the linear regression. That is, no significant effects of TMS group (i.e., stimulation site in first session) and no interaction between frame type and TMS group were found in the second session (Table A-4).

Based on these findings, we discarded the Post-TMS as baseline measurement of subjects' susceptibility to framing. In the following analyses, we therefore selectively considered the TMS session.

Table A-4: Confirmation of no effect of cTBS of DLPFC on framing susceptibility via logistic regression

	Variables	TMS session	Post-TMS session
Response	Left PFC	-0.062	-0.12
		(0.268)	(0.288)
	Right PFC	-0.0839	0.224
		(0.246)	(0.28)
	Frame	0.408***	0.254
		(0.152)	(0.186)
	Frame × Left PFC	0.057	0.199
		(0.242)	(0.23)
	Frame × Right PFC	0.0508	-0.156
		(0.228)	(0.196)
	Constant	-0.309	-0.421**
		(0.191)	(0.214)

Note: Logistic regression of responses (L, S) in target trials. Although responses differed significantly as a function of frame type, no effect of TMS or interaction between frame and TMS group was found in the TMS session. No significant effects were found in the Post-TMS session. Note: Coding scheme: S = 0, L = 1; Frame: GF = 0, LF = 1. Vertex group: baseline. Shown are regression coefficients and robust SE (in parentheses) clustered by subject. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In order to investigate whether a potential TMS effect on susceptibility to framing was obscured by different effects of the TMS stimulation on the four probability of winning levels, we investigated gambling rates in both LF and GF as a function of winning probability (for a summary of mean gambling rates per probability of winning level see Figure A-6). No significant probability of winning-dependent response differences between experimental groups were found by visual inspection of these plots. In order to statistically confirm this observation, we conducted a linear regression analysis of framing effects (i.e., $L_{LF} - L_{GF}$) for each probability of winning level with TMS groups as predictor variables. These analyses did not reveal any differences in gambling rates per probability of winning between groups in the TMS session (Table A-5)

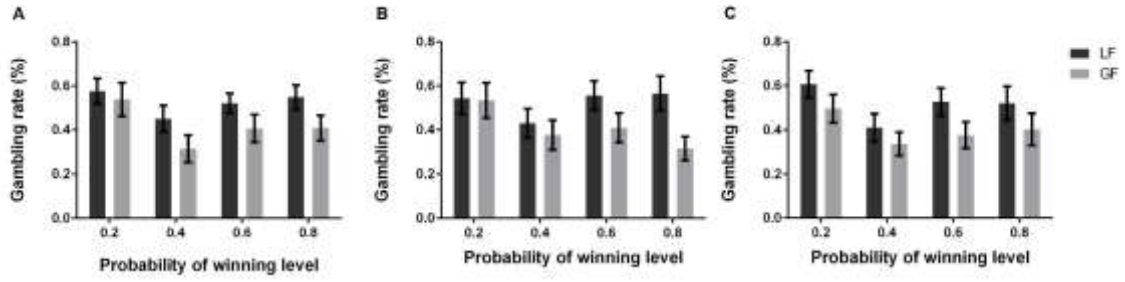

Figure A-6: Gambling rates per probability of winning level in LF and GF in (A) vertex group, (B) left PFC group and (C) right PFC group in the TMS session. Error bars: SEM

Table A-5: Influence of TMS on framing effect as a function of probability of winning level in the TMS session

	$(LF\text{-}GF)_{0.2}$	$(LF\text{-}GF)_{0.4}$	$(LF\text{-}GF)_{0.6}$	$(LF\text{-}GF)_{0.8}$
Right PFC	0.0717	-0.0353	0.0354	-0.0233
	(0.0749)	(0.0642)	(0.0746)	(0.09)
Left PFC	-0.029	-0.058	0.0323	0.111
	(0.0749)	(0.0642)	(0.0746)	(0.09)
Constant	0.0385	0.109**	0.115**	0.140**
	(0.0529)	(0.0454)	(0.0527)	(0.0637)

Note: We conducted linear regressions of framing effects, which were calculated as the difference in gambling rates between frames (L_{LF}-L_{GF}), on probability of winning level and TMS group in the TMS session. No group differences in the degree of framing for each probability level of winning were observed. Vertex group gamble rates served as baseline. Shown are regression coefficients and robust SE (in parentheses) clustered by subject. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In order to examine the finding of no TMS effect on the framing effect with a single-response sensitive method, we performed a logistic regression of responses (L, S) on probability of winning level, frame type and TMS group. Significant effects of probability of winning and interaction effects of frame type and probability of winning on gambling probability were observed. The probability of gambling was significantly decreased at a probability of winning level of 0.4 relative to 0.2 (baseline). Furthermore, the interaction between frame type and probability of winning level of 0.4 and 0.8 reflected framing effects at these probability levels. Most notably, no effects of TMS on participants' choices were found, nor did any TMS-relevant interaction reach significance (Table A-6). In other words, even when considering a potential differential effect of DLPFC-TMS on probability perception and associated interactions with frame type, we were not able to confirm a significant impact of disruption of DLPFC activity on susceptibility to framing.

Table A-6: No differential effects of cTBS on probability of winning perception in TMS session confirmed by means of logistic regression

	Variables	**TMS session**
Response	Left PFC	-0.0148
		(0.439)
	Right PFC	-0.166
		(0.392)
	Frame	0.154
		(0.2)
	Prob_0.4	-0.810***
		(0.143)
	Prob_0.6	-0.526*
		(0.29)
	Prob_0.8	-0.524
		(0.392)
	Left PFC × Prob_0.4	0.16
		(0.226)
	Left PFC × Prob_0.6	0.029
		(0.43)
	Left PFC × Prob_0.8	-0.395
		(0.605)
	Right PFC × Prob_0.4	0.135
		(0.229)
	Right PFC × Prob_0.6	0.0364
		(0.476)
	Right PFC × Prob_0.8	0.149
		(0.601)
	Frame × Prob_0.4	0.311**
		(0.157)
	Frame × Prob_0.6	0.31
		(0.252)
	Frame × Prob_0.8	0.416**
		(0.165)
	Left PFC × Frame	-0.12
		(0.303)
	Right PFC × Frame	0.291
		(0.284)
	Left PFC × Frame × Prob_0.4	-0.114
		(0.207)
	Left PFC × Frame × Prob_0.6	0.23
		(0.328)
	Left PFC × Frame × Prob_0.8	0.601
		(0.382)
	Right PFC × Frame × Prob_0.4	-0.422
		(0.244)
	Right PFC × Frame × Prob_0.6	-0.143
		(0.336)
	Right PFC × Frame × Prob_0.8	-0.391
		(0.296)
	Constant	0.152
		(0.303)

Note: Logistic regression of response (L, S) on TMS group (Left PFC, Right PFC, Vertex [baseline group]), frame type, dummy-coded probability of winning levels with 0.2 probability of winning as baseline and interactions between those terms. No effect of TMS group was found; neither did any TMS-relevant interaction reach significance. Note: Coding scheme as described before. Shown are regression coefficients and robust SE (in parentheses) clustered by subject. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

As a broad range of endowment amounts was used in the present study, we investigated whether TMS might have interacted with the endowment amount, thus leading to a change in framing effects between groups dependent on endowment level. For this reason, we performed a logistic regression of gambling probability comprising dummy variables for the TMS group and frame type as well as endowment amounts and interactions between those terms as predictor variables (Table A-7). For the TMS session, the significant effect of frame type on gambling behavior, i.e., the framing effect was again confirmed. No significant influences of TMS, endowment amount or significant interactions between these terms were observed. That is, neither did endowment amount per se lead to differences in susceptibility to framing, nor did we observe differential effects of TMS and interactions between those terms on susceptibility to framing.

Table A-7: No significant interaction effect between cTBS and endowment amount on gambling behavior

	Variables	**TMS session**
Response	Left PFC	-0.0504
		(0.32)
	Right PFC	-0.006
		(0.302)
	Frame	0.611***
		(0.235)
	Endowment	-0.001
		(0.002)
	Endowment × Frame	-0.0025
		(0.0022)
	Left PFC × Endowment	-0.00014
		(0.003)
	Right PFC × Endowment	-0.001
		(0.0027)
	Left PFC × Frame	-0.131
		(0.361)
	Right PFC × Frame	-0.274
		(0.311)
	Left PFC × Endowment × Frame	0.00234
		(0.004)
	Right PFC × Endowment × Frame	0.004
		(0.0028)
	Constant	-0.212
		(0.23)

Note: Logistic regression of responses (L, S) on the effect of TMS group, frame type and endowment level. A significant influence of frame type on gambling probability, that is, significantly increased probability of choosing L in LF relative to GF, was observed. However, neither endowment amount per se nor any interaction related to endowment and TMS group significantly influenced participants' choices. Note: Coding scheme as described above, endowment: original amounts used. Shown are regression coefficients and robust SE (in parentheses) clustered by subject. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

As previous studies revealed that personality traits influence individual susceptibility to framing (e.g., Levin et al., 2002), a further logistic regression model was set up in order to control for the influence of personality on the framing effect (summary statistics of questionnaires see Table A-8). It was observed that neuroticism was significantly associated with decreased probability of gambling. In addition, a negative relationship between conscientiousness and gambling probability was observed at trend level. However, controlling for the influence of different personality indices did also not reveal significant differences in framing susceptibility of the three TMS groups (Table A-9). In other words, even when we assumed that the role of DLPFC in framing susceptibility was obscured by personality traits and thus controlled for personality variables, we did not find differential effects of DLPFC-TMS on the framing effect.

Table A-8: Summary questionnaire scores

Questionnaire	Mean	SD
STAI-X1	32.72	5.77
STAI-X2	36.33	7.89
NC	32.8	23.03
NEO-N	18.59	7.79
NEO-E	29.6	6.89
NEO-O	32.95	6.71
NEO-A	31.84	5.91
NEO-C	30.71	6.8
BIS-11	61.44	8.47

Note: Shown are mean and standard deviation (SD). STAI-X1 and –X2: State-Trait-Anxiety Inventory, NC: Need for cognition scale; NEO-FFI subscales: N = neuroticism, E = extraversion, O = Openness, A = agreeableness, C = conscientiousness, BIS-11 = Barratt Impulsiveness Scale

Table A-9: Controlling for the influence of personality on the framing effect

	Variables											
Response	Frame	0.408***	0.409***	0.409***	0.408***	0.410***	0.409***	0.409***	0.408***	0.409***	0.461***	0.464***
		(0.152)	(0.153)	(0.153)	(0.152)	(0.153)	(0.153)	(0.153)	(0.153)	(0.153)	(0.154)	(0.155)
	Left PFC	-0.062	-0.0576	-0.0265	-0.0617	-0.123	-0.0867	-0.101	-0.0707	-0.0857	0.0129	-0.0247
		(0.268)	(0.25)	(0.275)	(0.269)	(0.262)	(0.268)	(0.272)	(0.267)	(0.275)	(0.264)	(0.242)
	Right PFC	-0.0839	-0.0826	-0.0643	-0.0842	-0.121	-0.0844	-0.107	-0.073	-0.105	-0.0362	-0.0455
		(0.246)	(0.248)	(0.245)	(0.255)	(0.24)	(0.247)	(0.25)	(0.247)	(0.248)	(0.239)	(0.24)
	Left PFC × Frame	0.057	0.061	0.0565	0.057	0.0564	0.0573	0.0575	0.0578	0.057	0.00489	0.0139
		(0.242)	(0.244)	(0.242)	(0.242)	(0.243)	(0.243)	(0.243)	(0.242)	(0.242)	(0.243)	(0.248)
	Right PFC × Frame	0.0508	0.0528	0.0503	0.0508	0.0512	0.0514	0.0515	0.0513	0.051	-0.000672	0.00259
		(0.228)	(0.229)	(0.228)	(0.228)	(0.229)	(0.229)	(0.228)	(0.228)	(0.228)	(0.229)	(0.231)
	NEO-N		-0.0192**									-0.0251***
			(0.00837)									(0.00961)
	NEO-E			-0.00837								-0.0266
				(0.0114)								(0.0165)
	NEO-O				-9.59^{-05}							0.00297
					(0.0122)							(0.0152)
	NEO-A					0.0224						0.0334**
						(0.0149)						(0.0156)
	NEO-C						-0.0170*					-0.0112
							(0.00925)					(0.0149)
	STAI-X1							-0.0142				-0.00454
								(0.0175)				(0.0221)
	STAI-X2								-0.00801			-0.000264
									(0.0114)			(0.0156)
	BIS-11									0.00737		0.00162
										(0.00969)		(0.0153)
	NC										-0.004	-0.00248
											(0.00299)	(0.00502)
	Constant	-0.309	0.0452	-0.0796	-0.306	-0.991**	0.222	0.177	-0.0188	-0.747	-0.249	0.215
		(0.191)	(0.253)	(0.366)	(0.455)	(0.493)	(0.333)	(0.632)	(0.483)	(0.596)	(0.205)	(1.627)

Note: For questionnaire abbreviations, see Table 7. Increased neuroticism scores lead to significantly decreased probability of lottery choices. Shown are regress coefficients and robust SE (in parentheses) clustered by subject. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Reaction time data

As cTBS may have affected cognitive processing speed in a more pronounced fashion than actual preferences, we investigated the effects of cTBS on reaction times (for mean reaction times, see Figure A-7). A linear regression on reaction times with TMS group, frame type and interaction between these terms as predictor variables identified a significantly faster mean reaction time in the GF relative to LF. No significant TMS group or interaction effects were observed in the TMS session (Table A-10).

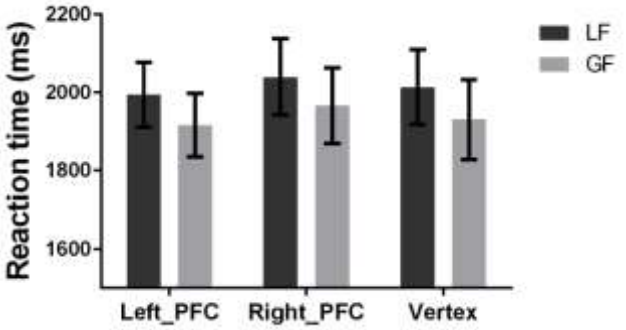

Figure A-7: Reaction times in LF and GF in the three treatment groups in the TMS session. Apparently, subjects took their decisions faster in the GF relative to LF. This pattern seemed to be true for all TMS groups.

Table A-10: Significantly faster reaction times in GF relative to LF, no effect of TMS on reaction time

Variables	**Reaction times (TMS session)**
Frame	83.41***
	(24.93)
Left PFC	-14.75
	(129.4)
Right PFC	35.82
	(139.7)
Frame × Left PFC	-4.804
	(41.03)
Frame × Right PFC	-9.712
	(36.87)
Constant	1931***
	(101.5)

Note: A linear regression analysis confirmed slower reaction times in LF relative to GF were observed. Note: Coding scheme: GF = 0, LF = 1. Vertex group: baseline. Shown are regression coefficients and robust SE (in parentheses). *** $p < 0.01$; ** $p < 0.05$¸* $p < 0.1$

In conclusion, we did not find evidence for the causal role of DLPFC in inhibition of the framing effect. cTBS of left or right DLPFC did not cause changes in preferences or response latencies since no TMS group differences of these parameters were found. Considering mean response data (via linear regression) as well as single responses (by means of logistic regression) did not reveal any differential effects of right or left DLPFC-TMS on susceptibility to framing. This absence of a TMS effect even held when taking into account potentially distinct effects of TMS on choice behavior dependent on probability perception, endowment amount and the moderating effects of personality.

A.5 Discussion

The current study investigated the causal role of the DLPFC in the framing effect, which describes a fundamental violation of the description invariance axiom of rational choice theory (Tversky & Kahneman, 1981, 1986). The DLPFC has been shown to be involved in the implementation of cognitive control of behavior in previous studies (MacDonald et al.,

2000; Miller & Cohen, 2001; Petrides, 2000). In particular, a recent neuroimaging study (De Martino et al., 2006) identified bilateral DLPFC as neural correlate of frame-incongruent decisions. De Martino et al. (2006) found significantly increased BOLD signal in trials, in which subjects were able to overcome the general behavioral tendencies imposed by the frames. Similar results were obtained by Gonzalez et al. (2005) who found increased DLPFC activation for risky relative to sure choices in the positive frame employing the classic Asian disease Problem (Tversky & Kahneman, 1986). Overcoming a behavioral heuristic seems to require additional cognitive resources, and the fMRI results suggested that this resistance to framing might necessitate recruitment of the DLPFC. Based on this correlational fMRI evidence for the involvement of the DLPFC in frame-incongruent choices (De Martino et al., 2006; Gonzalez et al., 2005) we hypothesized a causal role of DLPFC in inhibition of behavioral tendencies imposed by frames. Thus, the main purpose of the current study was to test this conjecture, by probing whether the DLPFC indeed is causally involved in inhibition of susceptibility to framing. cTBS was applied to the left or right DLPFC in the treatment groups and to the vertex in the control condition in order to investigate the effect of temporary DLPFC impairment on the framing effect. However, our DLPFC interference did not induce behavioral changes. That is, no effect of left or right DLPFC activity disruption on the susceptibility to framing was observed. Preferences between stimulation groups did not significantly differ from the control group in the TMS session. Logistic regression analyses revealed a stable framing effect across all three groups without significant group differences. cTBS did not lead to changes in probability of winning perception employed in the gambles. No interaction between cTBS and endowment amount was found. Controlling for various personality traits did also not reveal differential effects of unilateral cTBS of DLPFC on framing susceptibility. Furthermore, cTBS did not influence cognitive processing speed, as we did not find reaction time differences between DLPFC and vertex groups. Whereas increased reaction times were observed during the LF relative to GF across groups, no significant reaction time differences between groups were observed.

Taking into account these null results, we cannot draw a firm conclusion on the functional relevance of DLPFC for framing resistance. However, different hypothesis about the neural basis of resistance to framing and associated motivations for future studies can be drawn from our study.

First, based on the results by De Martino et al. (2006) it might be speculated that ACC - instead of DLPFC - is the brain structure with functional (i.e., causal) relevance for

framing resistance. In this scenario, computations in DLPFC might merely be correlated with ACC activation without any functional contribution to framing-related processing itself. In order to test this conjecture, we should apply cTBS to the ACC region identified by De Martino et. al (2006) followed by a framing task and observe whether a "virtual lesion" of ACC leads to changes in susceptibility to framing.

Second, we might hypothesize that framing susceptibility is regulated by synchronized activation of a network of brain regions (including the ACC as core component and DLPFC) involved in cognitive control processes. Transient disruption of one network component may be compensated by other functional components. That is, partial disruption of the putative ACC-DLPFC network might engage compensatory processes resulting in stable behavior, which might be an explanation for the current findings. Hence, we may speculate that unilateral DLPFC disruption is compensated by intact contralateral DLPFC and/or ACC activity. That is, bilateral disruption of DLPFC (for instance, via dual-site TMS (O'Shea, Taylor, & Rushworth, 2008)) or transcranial direct current stimulation (tDCS) electrode montage might be needed in order to disrupt the framing-related control function of DLPFC and induce behavioral changes. This hypothesis is supported by De Martino et al. (2006) who report that resistance against the frame is correlated with bilateral DLPFC activity. Furthermore, the simultaneous disruption of both DLPFC and ACC might be necessary in order to induce changes in framing susceptibility. Therefore, in order to distinguish the role of DLPFC from ACC in resistance to framing, dual-site TMS (O'Shea et al., 2008) might be required for simultaneous interruption of bilateral DLPFC as well as simultaneous inhibition of DLPFC and ACC. In this way, we may be able to investigate whether bilateral DLPFC is a mere coactivation of ACC-related framing resistance. Furthermore, functional and effective connectivity analyses in combination with concurrent TMS-fMRI may be particularly suited to test between the correlational versus causal role of DLPFC (and ACC) in framing susceptibility. This approach would be appropriate in order to identify framing-relevant neural connectivity patterns and TMS-induced changes on the functional and behavioral level, thus providing causal evidence for the neurobiological substrate of the framing effect and its inhibition. Importantly, the anatomical properties of the ACC require the use of a double-cone coil, which reduces the precision of the stimulation of the target area. However, previous studies have suggested that effective TMS of ACC may be made possible by means of this coil type (Hayward et al., 2007).

The purely behavioral Post-TMS session was conducted in order to yield a baseline measurement of the framing effect to calculate within-subject effects. Against our expectations, no consistent framing effect was found in the Post-TMS session. Although temporary stable framing effects were revealed in the behavioral pilot experiments we conducted, the temporal delay between the two framing tasks significantly differed between the pilot and actual TMS experiment. Whereas the two pilot sessions were separated by a break of approx. 15 minutes, during which subjects completed a questionnaire battery, TMS and Post-TMS session in the actual experiment were separated by approx. 14 days. Intermediate- to long-term learning and consolidation effects might have taken effect between TMS and Post-TMS session, diminishing the power of the framing manipulation. Future studies interested in within-subject effects should apply a different temporal schedule and align TMS and behavioral session closer to each other in order to avoid the repetition effects observed in the present study. Taking into account that the maximal disruptive effect of cTBS lasts for approximately 20 to 30 minutes (Huang et al., 2005), temporal delays of TMS and non-TMS session between 60 to 90 minutes seem feasible. Given that shorter delays between sessions seem to avoid repetition effects, future studies can randomize the order of TMS and behavioral session in order to provide full experimental randomization.

Further technical issues should find attention when considering our experimental results. We cannot rule out that our procedure was compromised by distance-dependent TMS effects. We cannot foreclose cTBS effects being weakened by the location of the scalp-distant DLPFC target, as rapid decline in magnetic field strength with distance of targeted cortical sites from scalp surface weakens actual TMS effects (Stokes et al., 2005). Although coil position was administered according to current technical guidelines (Rossi, Hallett, Rossini, Pascual-Leone, & Safety of TMS Consensus Group, 2009) and precisely controlled via our neuronavigation device, it is possible that changes in coil position relative to the gyrus may be necessary in order to optimize the magnitude of cortical stimulation. Coil position and head-coil angle may have to be changed to yield optimal cTBS application. Future studies may test such technical modifications to investigate whether DLPFC functioning can be more profoundly interrupted and whether this affects resistance to framing effects.

More generally, we have to point to the general issue concerning the interpretation of null results. Although we did not find evidence for the causal role of DLPFC in inhibition of the framing effect, we are not able to strongly conclude that DLPFC is not causally

involved in framing effect-related computations, as our experimental design might not have been sensitive enough to detect the effects of our experimental manipulation. For instance, it might be argued that in order to yield significant group differences, more subjects might be needed. However, exceedingly small effect sizes for the difference in framing effects between vertex and left or right TMS group ($d = 0.004$ and $d = 0.047$, respectively) were found, which seems to refute the power argument. In case of such weak effects, it is fair to nevertheless conclude that previous studies may have overestimated the role of the DLPFC in overcoming framing susceptibility, and that unilateral DLPFC is presumably not the main neural locus at which framing-related neural processes are causally implemented.

To recapitulate, the current TMS investigation did not reveal a causal role of unilateral DLPFC in inhibition of the framing effect as disruption of neither left nor right DLPFC led to significant changes in susceptibility to framing relative to the control condition. Further studies should employ methodological modifications of our experimental design described above in order to clarify the neurobiological mechanism underlying the generation and inhibition of the framing effect.

A.6 References

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for Cognition : eine Skala zur Erfassung von Freude und Engagement bei Denkaufgaben. *Zeitschrift Für Sozialpsychologie*, *Vol. 25, H. 2*, S. 147–154.

Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa und McCrae*. Göttingen: Hogrefe.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. doi:10.1037/0022-3514.42.1.116

Cassotti, M., Habib, M., Poirel, N., Aïte, A., Houdé, O., & Moutier, S. (2012). Positive emotional context eliminates the framing effect in decision-making. *Emotion (Washington, D.C.)*, *12*(5), 926–931. doi:10.1037/a0026788

Coutlee, C. G., & Huettel, S. A. (2012). The functional neuroanatomy of decision making: prefrontal control of thought and action. *Brain Research*, *1428*, 3–12. doi:10.1016/j.brainres.2011.05.053

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science (New York, N.Y.)*, *313*(5787), 684–687. doi:10.1126/science.1128356

De Neys, W. (2006). Dual processing in reasoning: two systems but one reasoner. *Psychological Science*, *17*(5), 428–433. doi:10.1111/j.1467-9280.2006.01723.x

Evans, J. S. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459. doi:10.1016/j.tics.2003.08.012

Fecteau, S., Knoch, D., Fregni, F., Sultani, N., Boggio, P., & Pascual-Leone, A. (2007). Diminishing risk-taking behavior by modulating activity in the prefrontal cortex: a direct current stimulation study. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(46), 12500–12505. doi:10.1523/JNEUROSCI.3283-07.2007

Figner, B., Knoch, D., Johnson, E. J., Krosch, A. R., Lisanby, S. H., Fehr, E., & Weber, E. U. (2010). Lateral prefrontal cortex and self-control in intertemporal choice. *Nature Neuroscience*, *13*(5), 538–539. doi:10.1038/nn.2516

Gonzalez, C., Dana, J., Koshino, H., & Just, M. (2005). The framing effect and risky decisions: Examining cognitive functions with fMRI. *Journal of Economic Psychology*, *26*(1), 1–20. doi:10.1016/j.joep.2004.08.004

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science (New York, N.Y.)*, *324*(5927), 646–648. doi:10.1126/science.1168450

Hayward, G., Mehta, M. A., Harmer, C., Spinks, T. J., Grasby, P. M., & Goodwin, G. M. (2007). Exploring the physiological effects of double-cone coil TMS over the medial frontal cortex on the anterior cingulate cortex: an H2(15)O PET study. *The European Journal of Neuroscience*, *25*(7), 2224–2233. doi:10.1111/j.1460-9568.2007.05430.x

Huang, Y.-Z., Edwards, M. J., Rounis, E., Bhatia, K. P., & Rothwell, J. C. (2005). Theta Burst Stimulation of the Human Motor Cortex. *Neuron*, *45*(2), 201–206. doi:10.1016/j.neuron.2004.12.033

Huang, Y.-Z., Rothwell, J. C., Lu, C.-S., Wang, J., Weng, Y.-H., Lai, S.-C., … Chen, R.-S. (2009). The effect of continuous theta burst stimulation over premotor cortex on circuits in primary motor cortex and spinal cord. *Clinical Neurophysiology*, *120*(4), 796–801. doi:10.1016/j.clinph.2009.01.003

Kahneman, D., & Frederick, S. (2007). Frames and brains: elicitation and control of response tendencies. *Trends in Cognitive Sciences*, *11*(2), 45–46. doi:10.1016/j.tics.2006.11.007

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263. doi:10.2307/1914185

Katayama, T., & Rothwell, J. C. (2007). Modulation of somatosensory evoked potentials using transcranial magnetic intermittent theta burst stimulation. *Clinical Neurophysiology*, *118*(11), 2506–2511. doi:10.1016/j.clinph.2007.08.011

Knoch, D., Gianotti, L. R. R., Pascual-Leone, A., Treyer, V., Regard, M., Hohmann, M., & Brugger, P. (2006). Disruption of right prefrontal cortex by low-frequency repetitive transcranial magnetic stimulation induces risk-taking behavior. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(24), 6469–6472. doi:10.1523/JNEUROSCI.0804-06.2006

Levin, I. P., Gaeth, G. J., Schreiber, J., & Lauriola, M. (2002). A New Look at Framing Effects: Distribution of Effect Sizes, Individual Differences, and Independence of Types of Effects. *Organizational Behavior and Human Decision Processes*, *88*(1), 411–429. doi:10.1006/obhd.2001.2983

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex in Cognitive Control. *Science*, *288*(5472), 1835–1838. doi:10.1126/science.288.5472.1835

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. doi:10.1146/annurev.neuro.24.1.167

O'Shea, J., Taylor, P. C. J., & Rushworth, M. F. S. (2008). Imaging causal interactions during sensorimotor processing. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *44*(5), 598–608. doi:10.1016/j.cortex.2007.08.012

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*(6), 768–774.

Petrides, M. (2000). 6 - Mapping Prefrontal Cortical Systems for the Control of Cognition. In A. W. Toga & J. C. Mazziotta (Eds.), *Brain Mapping: The Systems* (pp. 159–176). San Diego: Academic Press. Retrieved from http://www.sciencedirect.com/science/article/pii/B9780126925456500088

Rossi, S., Hallett, M., Rossini, P. M., Pascual-Leone, A., & Safety of TMS Consensus Group. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *120*(12), 2008–2039. doi:10.1016/j.clinph.2009.08.016

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3–22. doi:10.1037/0033-2909.119.1.3

Smith, S. M., & Levin, I. P. (1996). Need for Cognition and Choice Framing Effects. *Journal of Behavioral Decision Making*, *9*(4), 283–290. doi:10.1002/(SICI)1099-0771(199612)9:4<283::AID-BDM241>3.0.CO;2-7

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Stokes, M. G., Chambers, C. D., Gould, I. C., Henderson, T. R., Janko, N. E., Allen, N. B., & Mattingley, J. B. (2005). Simple Metric For Scaling Motor Threshold Based on Scalp-Cortex Distance: Application to Studies Using Transcranial Magnetic Stimulation. *Journal of Neurophysiology*, *94*(6), 4520–4527. doi:10.1152/jn.00067.2005

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. doi:10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science (New York, N.Y.)*, *211*(4481), 453–458.

Tversky, A., & Kahneman, D. (1986). Rational Choice and the Framing of Decisions. *The Journal of Business*, *59*(4), S251–S278. doi:10.2307/2352759

Appendix

B Anticipatory anxiety disrupts neural valuation during risky choice

Jan B. Engelmann[†1,2], Friederike Meyer[†1], Ernst Fehr[+1] & Christian C. Ruff[+1]

[1]Laboratory for Social and Neural Systems Research (SNS-Lab), Department of Economics, University of Zurich, Zurich, Switzerland

[2]Donders Institute for Brain, Cognition and Behaviour, Centre for Cognitive Neuroimaging, Radboud University, Nijmegen, The Netherlands

[†] shared first authorship

[+] shared last authorship

B.1 Abstract

Incidental negative emotions unrelated to the current task - such as background anxiety - can strongly influence decisions. This is most evident in psychiatric disorders associated with generalized emotional disturbances. However, the neural mechanisms by which incidental emotions may affect choices remain poorly understood. Here we study the effects of incidental anxiety on human risky decision-making, focusing on both behavioral preferences and their underlying neural processes. Although observable choices remained stable across affective contexts with high and low incidental anxiety, we found a clear change in neural valuation signals: During high incidental anxiety, activity in ventromedial prefrontal cortex and ventral striatum showed a marked reduction in (1) neural coding of the expected subjective value (ESV) of risky options, (2) prediction of observed choices, (3) functional coupling with other areas of the valuation system, and (4) baseline activity. At the same time, activity in the anterior insula showed an increase in coding the negative ESV of risky lotteries, and this neural activity predicted whether the risky lotteries would be rejected. This pattern of results suggests that incidental anxiety can shift the focus of neural valuation from possible positive consequences to anticipated negative consequences of choice options. Moreover, our findings show that these changes in neural value coding can occur in the absence of changes in overt behavior. This suggest a possible pathway by which background anxiety may lead to the development of chronic reward desensitization and a maladaptive focus on negative cognitions, as prevalent in affective and anxiety disorders.

B.2 Introduction

Behavior is strongly guided by integral emotions that provide affective information about which behavioral options should be pursued or avoided (Rick and Loewenstein, 2008). How such goal-directed valuation of choice options is instantiated in the human brain is increasingly understood (Rangel et al., 2008; Lee et al., 2012; Levallois et al., 2012; Rushworth et al., 2012) and it is commonly assumed that dysfunctions of these goal-directed neural valuation processes may account for behavioral pathologies in psychiatric diseases (Sharp et al., 2012). However, emotions can also guide behavior more indirectly,

for instance when incidental background emotions (e.g., general or dispositional anxiety) influence decisions despite not being relevant for the task at hand. Such effects of incidental emotions on decision-making can be strongly detrimental, as evident in psychiatric disorders such as anxiety disorders (Maner et al., 2007; Giorgetta et al., 2012; Grecucci et al., 2013), depression (Harle et al., 2010; Engelmann et al., 2013), or mania (Minassian et al., 2004). Despite their obvious importance, not much is known about the neural mechanisms by which incidental emotions such as anxiety can affect behavioral choices (Knutson et al., 2008; Harlé et al., 2012).

The lack of knowledge about neural mechanisms for interactions of incidental emotions and choices appears somewhat surprising, given the substantial overlap in the neural circuits involved in decision-making and the processing of affective states. For instance, most core structures of the brain's putative valuation network - encompassing the VMPFC, VS, insula and amygdala (Hsu et al., 2005; De Martino et al., 2006; Preuschoff et al., 2008; Levy and Glimcher, 2012) - are also substantially involved in processing anxiety (Hartley and Phelps, 2012). It is therefore likely that incidental affective and task-related signals computed in these same areas may interact, thereby providing a neural substrate for immediate influences of background emotions on neural valuation during choice. Moreover, persisting incidental affective states may shape processing in these neural structures via neural plasticity, thereby possibly leading to chronic influences on neural computations involved in value-based choice (Paulus and Yu, 2012; Huys et al., 2013).

Here, we investigate with behavioral, electrophysiological, and neural data how experimentally-induced anticipatory anxiety impacts on risky decision-making. We chose to induce anticipatory anxiety with a threat-of-shock technique (Schmitz and Grillon, 2012) that has been successfully used in previous studies (e.g., Robinson et al., 2011). In this approach, anticipatory anxiety is induced as a physiological and affective response to an ongoing threat of painful stimulation that is temporally unpredictable and incidental to the task (Robinson et al., 2013b). Using this paradigm during fMRI and electrophysiological recording, we thus manipulated our participants' affective state while they completed a set of risky choices involving monetary gains and losses. This allowed us to investigate effects of anticipatory anxiety on individual decision-making and underlying neural activity. Importantly, due to our within-subject design, we could adequately control for potentially confounding variables known to influence risk attitude, such as sociodemographic, trait and genetic factors (Dreher et al., 2009; Capra et al., 2013).

We find that incidental anxiety can fundamentally change the mechanisms of value-coding during risky choice in the absence of overt behavioral changes. During anticipatory anxiety, neural coding of subjective value in VMPFC and VS was decreased but neural coding of negative subjective value in anterior insula was increased, thereby suggesting a shift in value-coding from predicted positive consequences toward a focus on possible negative outcomes. This change in neural value coding was clearly linked to observed behavior, as choices during safe trials could be predicted from BOLD signals in the VMPFC and VS (but not insula) whereas choices under threat could be predicted from the insula (but not VMPFC and VS).

B.3 Materials and methods

Subjects

43 right-handed male subjects (mean age [SEM] = 22.56 [1.42]) without a history of psychiatric or neurologic disorder participated in the study. In total, 10 subjects had to be excluded due to technical and image acquisition problems (excessive head movement: n = 1, scanner crash = 3, sequence problem: n = 1) or non-compliance with the task (n = 5, estimation of behavioral models did not converge due to choice patterns that showed too little consistency, which is a common observation (Sokol-Hessner et al., 2009). This left a total of 33 subjects that were included in all subsequent analyses. Written informed consent was obtained from all subjects and all procedures were approved by the ethics committee of the canton of Zurich, Switzerland.

Experimental timeline

Approximately one week before the experiment, participants were invited to the lab for an "endowment" session in which they were paid for completing a psychological questionnaire battery chosen to assess a comprehensive range of personality characteristics. This battery comprised the DOSPERT (Weber et al., 2002) to assess domain specific risk attitude, the BIS-11 (Patton et al., 1995) to assess impulsivity, the NEO-FFI (Borkenau and Ostendorf, 1993) to measure personality characteristics, the STAI (Spielberger et al., 1983) to assess anxiety symptoms and the BDI (Hautzinger et al., 1995) to measure depressive symptom severity. The participation fee for this session was a CHF 50 (ca. US$ 58 at time of testing) credit voucher. Importantly, it was stressed to the subjects that the money earned during the

endowment session would be their starting balance for the economic games played out in the second session, to avoid excessive risk-taking in that session due to the well-documented house-money effect (Thaler and Johnson, 1990).

During the second session, fMRI and electrodermal activity were recorded during a risky-decision making task (see section "Decision-making task"). The task was preceded by a short training session and an individual stimulation thresholding procedure (see section "Emotion induction technique"). The whole session lasted a total of 120 minutes (including subject preparation, scanning time and post-experimental debriefing) and was remunerated with an additional show-up fee of CHF 50.

Experimental Procedures and Task

Emotion induction technique. Anticipatory anxiety was induced by giving participants advance information about impending but temporally unpredictable electric stimulation in two distinct contexts. In the threatening context ("threat trials" triggering anticipatory anxiety), the stimulation was painful, whereas in the matched safe context ("safe trials"), the stimulation was just noticeable but not painful. To this end, we administered electrical stimulation with two DS5 Isolated Bipolar Constant Current Stimulators (Digitimer Ltd., Welwyn Garden City, UK) and custom-made fMRI-compatible 5 mm ring electrodes. Two of these electrodes were positioned above the first and fourth dorsal interossei muscles of the left dorsal hand, respectively, to provide two separate locations for strong and weak stimulation intensities. This was implemented to avoid desensitization or carry-over effects between both contexts. The intensities employed for stimulation were individually calibrated before the scanning session with a well-established procedure (e.g., Singer et al., 2004). For this purpose, electrical pulses of different intensities were repeatedly given in a randomized order and participants rated each pulse on a visual analogue scale ranging from 0 (not painful at all/hardly perceptible) to 10 (unbearably painful) in steps of 1. Stimulus intensities corresponding to an individual rating of 1 and 8 were chosen for the weak versus strong experimental treatments, respectively. In order to control for sensitization or desensitization, subjects had to rerate both low- and high-stimulation intensity after each fMRI run and stimulus intensity was set for each run to the rerated values of 1 and 8, respectively.

During each fMRI run, stimulation could occur at unpredictable time points during decision-making periods, thus creating threat or safe periods during which participants were expecting strong versus weak electrical stimulation, respectively. The two affective contexts

were presented in a blocked fashion. During each block, participants made four risky decisions as outlined in detail below. In order to augment the efficacy of anticipatory anxiety induction, the number and time points of electrical stimulation events throughout both threat and safe blocks were determined to be completely unpredictable for subjects. For this purpose, the number of stimulation events was determined for each block by random draw from a gamma distribution (shape parameter = 1; scale parameter = 1). The exact timing of these stimulation events was then determined at random time points between the offset of the cue display and onset of the resting screen, as determined by drawing from a uniform distribution with the constraint that successive electrical shocks were separated by at least 0.2 s.

The emotion induction was conducted in a within-subject design: Each subject took decisions in both affective contexts and therefore could serve as his own control. This has significant advantages over a between-subject design, where the effects of affective context could be confounded by individual differences in variables that are known to influence risk attitude, such as sociodemographic, personality and genetic factors (Dreher et al., 2009; Capra et al., 2013). On the flip side, the within-subject design chosen here comes with the small risk that participants may make their choices based on memory when confronted with a specific lottery for a second time, in a different affective context. This conjecture appears highly unlikely, as subjects would need to remember their choices for 280 trials that were presented in random order. We nevertheless took care to inspect our data for such effects, by conducting regression analyses to test whether reaction times were different for lotteries that were presented for the first or the second time (such differences would be expected if participants responded from memory, which requires less deliberation). No such effects were revealed by these analyses (see section "Results"), suggesting that participants employed the same strategies to evaluate the first and second presentation of each lottery.
All visual and tactile stimuli were presented and recorded with Cogent 2000 (http://www.vislab.ucl.ac.uk/cogent.php) implemented in MATLAB (The MathWorks Inc, Natick, Massachusetts).

Decision-making task. The fMRI experiment consisted of 7 runs (each with 472 seconds average duration) during which participants made risky decisions in the threat and safe contexts. Each of these runs (see Figure B-1a, timeline) comprised 5 threat blocks with strong stimulation intensity and 5 safe blocks with weak stimulation intensity. These blocks lasted 43 seconds on average and were pseudorandomly interleaved so that not more than 2 blocks of the same affective valence would follow each other. The single experimental

blocks were interleaved with resting blocks of 6 seconds duration to allow sustained hemodynamic activity to return to baseline.

During each block, 4 decision-making trials were presented to the participants. On each trial, participants had to choose between taking part in a lottery with two equiprobable (50/50) options and a sure outcome. In 5 of the 7 functional runs, we employed *mixed gambles* in which the lottery led with 50% probability to a win of amount x_1 and with 50% probability to a loss of amount x_2. The sure outcome was always a zero change in income (see Figure B-1b, gamble composition). The relationship between gain and loss amounts was chosen to encourage a reasonable amount of risk taking. Previous research indicates that participants overweight losses relative to gains in a multiplicative fashion (Kahneman and Tversky, 1979). We therefore employed prospect theory to parameterize general risk-taking and this overweighting (see section "Behavioral data analysis") and to inform the analysis of neural valuation processes. All imaging analyses were performed based on the 5 mixed gamble runs.

In order to facilitate parameter estimation for the behavioral choice models, we also included 2 additional functional runs composed of gain-only and loss-only gambles. These "pure gamble" runs were presented in random positions within the sequence of mixed-gamble runs and included an equal number of gain only (20) and loss only trials (20) within each run. Pure gambles differed from mixed gambles in that lottery outcomes were now either both positive (50% probability to *win* amount x_1 or x_2) or both negative (50% probability to *lose* amount x_1 or x_2). Lottery amounts in the gain-only domain varied between 0 and 30 (in steps of 10) for x_1 and 20 and 45 (in steps of 5) for x_2. Lotteries in the loss-only domain varied between 0 and -30 for x_1 (in steps of 10) and -20 and -45 (in steps of 5) for x_2. The sure outcome was always a non-zero amount generated by calculating a certainty equivalent corresponding to different levels of risk aversion for each pure lottery. In order to capture different degrees of risk aversion, five different degrees of risk aversion were assumed for both the gain domain and loss domain; each of these parameter values was employed for the calculation of the certainty equivalent for four different lotteries in both the threat and safe context. The certainty equivalent (CE) was calculated as follows:

(Equation 1) $$CE_{Gain} = (0.5\ x_1^{\alpha} + 0.5x_2^{\alpha})^{1/\alpha}$$ and

(Equation 2) $$CE_{Loss} = (0.5\ x_1^{\beta} + 0.5x_2^{\beta})^{1/\beta}$$

where α and β capture the degree of risk aversion in the gain domain and loss domain of the value function, respectively. The following values of α and β were used in order to create

CEs: $\alpha = \beta = [0.2, 0.5, 0.7, 0.9, 1.3]$ (see detailed description in the section "Behavioral data analysis").

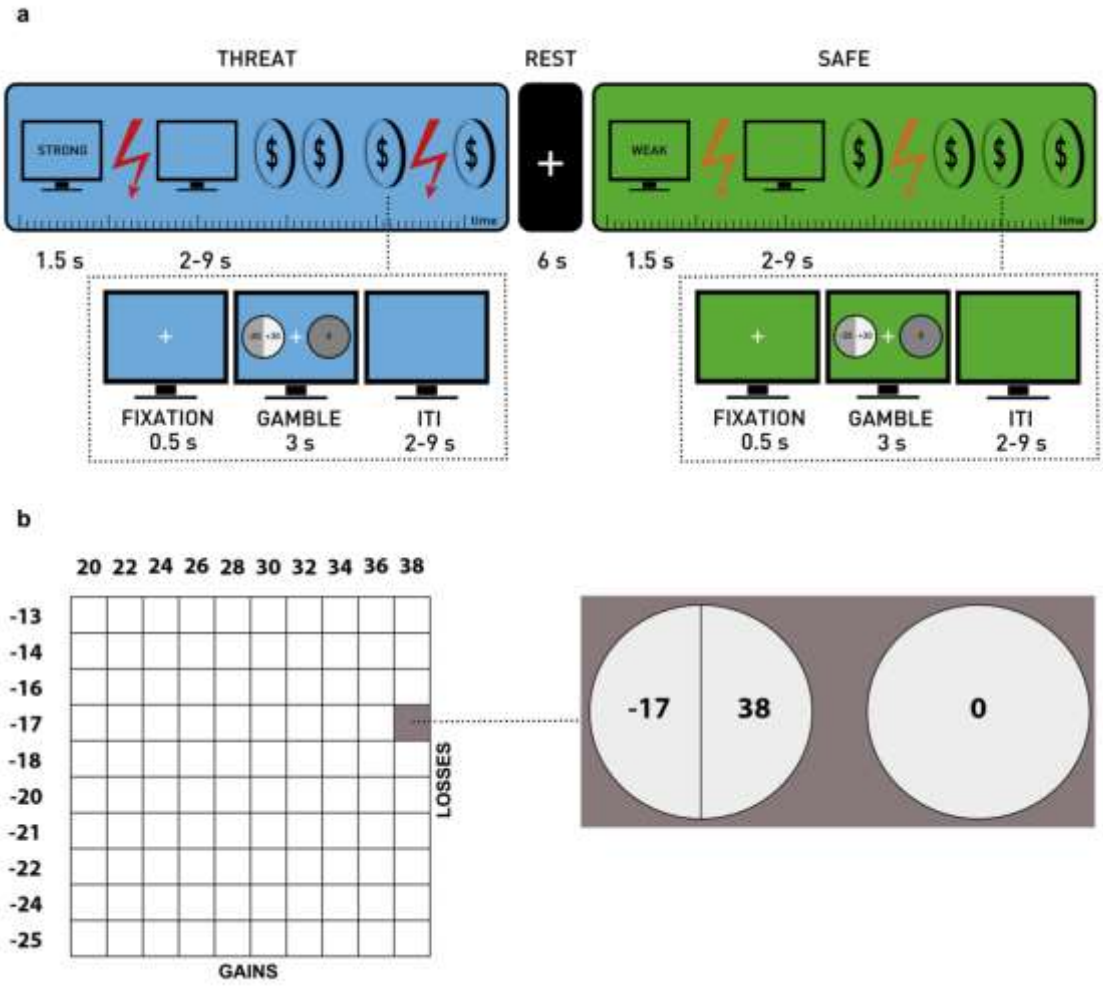

Figure B-1: Hybrid fMRI-Design. (a) The top row illustrates the timing of different blocks and the bottom row the timing of the gambles. Blocks that were threatening (strong stimulation) or safe (weak stimulation) were randomly interleaved (two of ten blocks per run are shown). The type of block was indicated by the screen color (blue or green), which was constant throughout the block. At the beginning of each block, a cue (STRONG or WEAK) was displayed, followed by delivery of a "reminder shock" (dark or light red arrow, respectively, following the cue display). After a jittered ITI (empty screen), the beginning of a gamble was signaled by a fixation cross, followed by the gamble screen. Subjects were instructed to choose their preferred option (lottery or sure outcome) as fast and accurately as possible via button press. Each block comprised 4 gambles and was followed by a REST period. Electrical stimulation was delivered randomly and unpredictably during the whole block (indicated by the varying positions of dark red [strong stimulation] or light red [weak stimulation] arrows, the only stimulation-free phases were the rest periods). (b): Possible gains and losses for mixed gambles. Gain and loss values were sampled from the gain-loss matrix depicted on the left (columns: gains; rows: losses). Each of the resulting combinations was shown once in the safe and once in the threat condition. For this purpose, the lottery with the corresponding outcomes was shown on one side of the screen, whereas the corresponding safe outcome was shown on the other (see right panel for example screen). The presentation sides for lottery option and sure option outcomes were counterbalanced across participants.

Payment schedule. Seven choices were randomly drawn and played out after the experiment (one per functional run). Their mean outcome (which could be negative or positive) was added to the CHF 100 show-up fee for both sessions in order to determine the participants' final payoff.

fMRI data acquisition. Functional magnetic resonance images (fMRI) were collected using a 3T Philips Achieva whole-body magnetic resonance scanner (Philips Medical Systems, Best, The Netherlands) equipped with an 8-channel Philips sensitivity-encoded (SENSE) head coil. Structural image acquisition consisted of 180 T1-weighted transversal images (0.75 mm slice thickness). For functional imaging, T2* images were obtained using a SENSE T2*-weighted echo-planar imaging sequence (Pruessmann et al., 1999) with an acceleration factor of 2.0. The sequence consisted of 33 axial slices covering the whole brain (slice thickness = 3mm, inter-slice gap = 0.75 mm, ascending sequential acquisition, flip angle = 78°, repetition time = 1750 ms, echo time = 30 ms, field of view = 240 mm, matrix size = 80 × 80). To optimize functional sensitivity in orbitofrontal cortex and the medial temporal lobes, we used a tilted acquisition in an oblique orientation at -20° relative to the AC-PC line (Deichmann et al., 2003).

Psychophysiological measures. In order to confirm efficacy of the threat-of-shock treatment, we acquired skin conductance responses (SCRs) using a PowerLab 4/25T amplifier with a GSR Amp (ML116) unit and a pair of MR-compatible finger electrodes (MLT117F). The electrodes were attached to the participants' left index and ring finger using conducting gel and dedicated Velcro. Recordings were performed with Lab Chart 5 software, with the recording range set to 40 µS and using initial baseline correction ("subject zeroing") to subtract the participant's absolute level of electrodermal activity from all recordings (devices and software from ADInstruments Pty, Ltd., Bella Vista, Australia). Each participant's SCRs were initially smoothed with a running average over 500 samples (0.5 s) to reduce scanner-induced noise. Data were then resampled from 1 KHz to 1 Hz and subsequently z-transformed. In line with previously established methods (Bach et al., 2009; Choi et al., 2012), we estimated mean SCRs during the decision periods of both affective conditions with multiple linear regression analysis, as implemented in Analysis of Functional NeuroImages (AFNI; Cox, 1996). The statistical model contained a total of 6 regressors that reflected onset times of choices under expectancy of strong and weak electrical shocks, cue times indicating the onset of a block, and delivery times of strong and weak electrical shocks. Average responses to each trial type were estimated via deconvolution from event onset to 16 s post onset using 17 cubic spline basis functions.

Baseline drifts of the SCR were modeled with constant linear and quadratic terms included for each run. For these analyses, SCR data sets of two subjects had to be excluded due to acquisition problems.

Behavioral data analysis. Behavioral data were analyzed using R (www.r-project.org) and Stata (StataCorp LP, College Station, Texas, USA). As a first step, reaction times (RTs) and gambling rates (percentage of trials where lotteries were chosen) were compared between the threat and safe contexts. Moreover, we inspected the data for adaptation and order effects by means of a regression model (with robust standard errors clustered by subject). This model regressed trial-wise reaction times on variables representing each choice's presentation order (first versus second presentation in the experiment), affective context (threat vs safe), gain and loss amount, the experimental trial (1 to 280, to account for learning effects), and all two-way interactions of these variables.

In a second step, Cumulative Prospect Theory (CPT; e.g., Bruhin et al., 2010) was used to derive subject-specific prospective utilities U(x) of each lottery x with outcomes x_1 and x_2. The lottery in each trial was formalized as:

$$\text{(Equation 3)} \qquad x = (x_1, p_1; x_2, p_2)$$

where p_i is the probability of obtaining outcome x_i, $i \in \{1, 2\}$, and $p_1+p_2=1$. In order to calculate subject-specific values of the lottery, we fitted the value function to of prospect theory determine the value $v(x_i)$ for each outcome x_i:

$$\text{(Equation 4)} \quad v(x_i) = \begin{cases} x_i^{\alpha} & \text{if } x_i \geq 0 \\ -\lambda\,(-x_i)^{\beta} & \text{if } x_i < 0 \end{cases}$$

where α quantifies risk aversion in the win-domain, β quantifies risk aversion in the loss-domain, and λ quantifies overweighting of losses relative to gains. We allowed risk preferences to vary between the threat and safe conditions by defining:

$$\text{(Equation 5)} \qquad \alpha = (1 - \tau) \times \alpha_S + \tau \times \alpha_T$$

$$\text{(Equation 6)} \qquad \beta = (1 - \tau) \times \beta_S + \tau \times \beta_T$$

$$\text{(Equation 7)} \qquad \lambda = (1 - \tau) \times \lambda_S + \tau \times \lambda_T$$

where

$$\text{(Equation 8)} \quad \tau = \begin{cases} 1 & \text{if threat trial} \\ 0 & \text{if safe trial} \end{cases}$$

and subscript S denotes parameters for safe trials, whereas subscript T denotes parameters during threat trials. The prospective utility of a gamble, U(x), is defined as

$$\text{(Equation 9)} \qquad U(x) = [p_1 \times v(x_1)] + [p_2 \times v(x_2)]$$

where we ignore the nonlinear probability weighting typically present in prospect theory models because we restricted $p_1 = p_2 = 0.5$.

To determine the probability p of choosing the lottery (LO) over the sure option (SO), we employed a form of the logit probabilistic choice rule that allows for noise in option selection via the free parameter μ:

$$\text{(Equation 10)} \qquad p\,(LO, SO) = 1/[1+ (e^{-\mu(U(LO) - U(SO))})]$$

fMRI data analysis. Preprocessing and statistical analyses of functional data were performed using SPM8 (Wellcome Department of Imaging Neuroscience, London, UK). All functional volumes were realigned to the first volume using b-spline interpolation and subsequently unwarped using fieldmaps estimated by SPM to remove residual movement-related variance due to susceptibility-by-movement interactions (Andersson et al., 2001). To improve coregistration, bias-corrected and coregistered anatomical and mean EPI images were created using the New Segment toolbox in SPM. The forward deformation fields created in the context of the nonlinear normalization of individual gray matter tissue probability maps were then employed to normalize the functional timeseries to the Montreal Neurological Institute (MNI) T1 template. Finally, functional data underwent spatial smoothing using an isotropic 6-mm FWHM Gaussian kernel.

Statistical analyses were carried out using the general linear model (Friston et al., 1994). Regressors of interest were modeled using canonical hemodynamic response functions temporally aligned with the onset of the events of interest. Time and dispersion derivatives were added in order to account for subject-to-subject and voxel-to-voxel variation in response peak and dispersion (Henson et al., 2002). Neural processes triggered by lottery presentation per se (averaging over all different lotteries) were modeled with two

regressors (trials in safe context, trials in threat context). Neural valuation processes on each trial were modeled with two parametric regressors (one for threat trials and one for safe trials) that coded each lottery's neural expected subjective value (ESV; e.g. Levy et al., 2010), as also done in other studies of mixed gambles (Minati et al., 2012a, 2012b; Sokol-Hessner et al., 2013). ESV was defined for each gamble as the neural representation of the behaviorally measured U(x) (see section "Behavioral data analysis") and was assumed to linearly correlate with U(x) (Levy et al., 2010). As the SO was 0 for all mixed gamble trials, it was omitted from the GLM. The value regressors in the GLM spanned for each trial the decision period from the onset of the decision screen until the subject's choice, as indicated via button press. Estimation of the behavioral model revealed that the loss aversion coefficients λ_T and λ_S were not significantly different from each other. In a similar vein, no significant deviations from 1 were found for the risk aversion parameters α_S, α_T, β_S, and β_T (see section "Results"). We therefore employed the individual mean loss aversion coefficient $\lambda = (\lambda_T + \lambda_S)/2$ as subject-specific input to derive the predicted neural valuation processes for each trial. Note, however, that inclusion of the estimated loss aversion coefficients in the T and S condition, λ_T and λ_S, instead of the average of the two conditions, did not lead to significantly different fMRI results. Additional regressors added to the first-level GLM corresponded to the cues at the beginning of the blocks, actual electrical stimulation at weak and strong intensities, and trials where participants omitted responses. Main effects and interactions of affective context and decision-making were computed with linear contrasts of the first-level parameter estimates, resulting in one contrast estimate for each of these effects in each participant.

Statistical inference was performed with second-level random-effects comparisons of the first-level contrasts representing the trial onset and ESV regressors for the threat and safe context. We tested our hypotheses about changes in neural valuation due to anticipatory anxiety in regions of interest (ROIs) previously shown to be involved in neural valuation and emotion processing. These regions comprised VMPFC (Levy and Glimcher, 2011; Winecoff et al., 2013), VS (Tom et al., 2007; Bartra et al., 2013; Robinson et al., 2013a), insula (Robinson et al., 2013a), and amygdala (Jenison et al., 2011; Fossati, 2012). Bilateral masks for amygdala and insula were taken from the WFU PickAtlas (http://fmri.wfubmc.edu/software/PickAtlas). For VMPFC and VS, which are anatomically less distinctively defined, we employed MNI coordinates and extent estimations from the published literature. Coordinates for the VMPFC mask (right VMPFC: x = 4.27, y = 35.18, z = 11.82; left VMPFC: x = -7.29, y = 38, z = -10.57; spheres with 12 mm radius) were

taken from a meta-analysis of neural valuation processes (Levy and Glimcher, 2012). Bilateral coordinates for the ventral striatum (VS; bilateral inferior VS (nucleus accumbens): x = ±9, y = 9, z = -8; superior ventral striatum (ventral caudate): x = ±10, y = 15, z = 0; spheres with 6 mm radius) were taken from a study employing connectivity-based parcellation of the striatum (Di Martino et al., 2008), which has also been used for this purpose by other studies (Kelly et al., 2009). These eight regions were then combined into a simple ROI mask using WFU PickAtlas "Union" function. All analyses of value-related responses were restricted to this inclusive mask with the statistical threshold set at $p < 0.001$, uncorrected, in line with the threshold commonly used in *a priori* defined hypothesis tests of value-related responses in these regions (Plassmann et al., 2010; Kang et al., 2011; Litt et al., 2011; Lin et al., 2012; Sokol-Hessner et al., 2012). Additionally, we ran exploratory whole-brain analyses with a statistical threshold of $\boldsymbol{p} < 0.05$ family-wise error (FWE) corrected for the full brain.

Trial-by-trial regression analysis of the relationship between ROI signal and subjective value. To confirm the results of the SPM analyses with independent data, we examined the link between subjective values and neural activity in analyses of the trial-by-trial time course data extracted from the VMPFC, VS and insula. We employed the leave-one-subject-out (LOSO) cross-validation procedure (Esterman et al., 2010), in which regions for extraction of each subject's data are defined by a GLM analysis of value responses in the other N-1 subjects. This procedure ensures that ROI analyses are performed on new data that are independent of those used for ROI definition, thereby avoiding circular inference (Kriegeskorte et al., 2009) and allowing within-sample replication in an independent dataset.

Specifically, activation time courses were extracted from spherical volumes of interest (VOI) centered on the activation foci from each LOSO GLM within the VMPFC (12 mm radius), VS (6 mm radius) and insula (12 mm radius). Data were spatially averaged over the sphere and were temporally downsampled to one value for each trial, by averaging over 4 images within the period of 3.5 to 10.5 seconds post-trial onset. This ensured that the data summarized the delayed and dispersed hemodynamic response reflecting choice-relevant computations (RTs ranged from 0.48 to 6.16 seconds post-stimulus onset). Spatial and temporal averaging and correction for session and movement confounds was implemented using the VOI toolbox in SPM8.

As a first step, we examined whether trial-by-trial activity within our regions of interest encodes subjective value differentially during the presence of threat relative to no threat. To

this end, we regressed trial-by-trial signal changes in our ROIs on subjective value, threat, and their interaction, using the following model:

$$\text{(Equation 11) } Signal_{ti} = \beta1 * SV_{ti} + \beta2 * T_{ti} + \beta3 * (SV_{ti} * T_{ti}) + \beta4 * X_{ti} + e$$

where for each for subject *i* on trial *t*, Signal is the regional signal within a given ROI, SV reflects the subjective value, T is a dummy variable that encodes the presence of threat and X reflects a vector of subject-specific confounding variables that include age, trait anxiety, the emotional stability subscale of the NEO-FFI, and trial number to control for temporal variation in the signal.

Prediction of choices from activation time courses in VOIs. To demonstrate the link between brain activity in each region of interest and the overt economic choices, we performed prediction analyses of choices based on neural signals (Kuhnen and Knutson, 2005; Berns et al., 2008). These analyses were conducted to investigate whether neural signals from independently-defined regions in VMPFC, VS, and insula predict overt choice, and how this predictive relationship changes with affective context (threat vs. safe). To this end, we conducted logistic regression of mixed-gamble acceptance on neural signals (extracted from subject-specific regions using the identical LOSO approach as described above), affective context (threat, safe), and their interaction using the following model:

$$\text{(Equation 12) } Choice_{ti} = \beta1 * Signal_{ti} + \beta2 * T_{ti} + \beta3 * Signal_{ti} * T + \beta4 * X_{ti} + e$$

where for each subject *i* on trial *t*, Choice is a dummy variable that reflects the acceptance of mixed gambles, Signal is the regional signal within a given ROI, T is a dummy variable that codes the presence of threat, and X reflects a vector of subject-specific confounding variables that include age, trait anxiety, the emotional stability subscale of the NEO-FFI, and trial number to control for temporal variation in the signal.

For both regression models, we first conducted analyses of all trials to test how affective context interacts with subjective values (Equation 11; ordinary least squares) or regional BOLD signals (Equation 12; Logit). To further characterize these interactions, we then conducted follow-up analyses of only threat trials and only safe trials, using a simpler model that did not include the affective-context factor and its interaction. The parameters of all regression models were estimated using robust and clustered standard errors that correct for heteroscedasticity and correlated responses from each subject using the Huber-White method implemented in the robcov function of the Regression Modeling Strategies (RMS)

package in R. Statistical inference on the estimated parameters was performed at a one-tailed $p < 0.05$, given the clear hypotheses in these confirmatory analyses.

Psychophysiological interaction analysis. To test whether the induction of anticipatory anxiety caused changes in the functional connectivity within the valuation network (VMPFC, VS; insula, amygdala) and with regions outside this network, we conducted psychophysiological interaction (PPI) analyses using a generalized form of context-dependent psychophysiological interaction analysis (McLaren et al., 2012). PPI regressors were added to the same statistical model as described above (but omitting parametric modulators). Seed regions for PPI analysis were defined as spheres with 3 mm radius centered on peak voxels of activation found in the ESV-related contrasts of the GLM analysis. To obtain an estimate of neural activity within the seed region, the first temporal eigenvariate of a principle component analysis (PCA) of the BOLD signal in all voxels of the seed region was extracted and deconvolved (Gitelman et al., 2003). A new GLM was then estimated for each subject that included the original design matrix and the following additional regressors: the deconvolved time course from the seed region and two psychological interaction regressors (risky decisions in the threat condition and the safe condition). These regressors were created by multiplying the neural activity in the seed region during the relevant decisions with the regressors corresponding to the decision-related activity in the two contexts (condition-specific on- and offset times convolved with the canonical HRF). PPI contrast maps for each subject were entered into a second level random-effects group analysis using the same inference procedures as for the GLM analyses described above.

Please note that these PPI analyses were conducted on the residuals of the original GLM model; this ensures that any changes in functional coupling are estimated after the average activity elicited by each decision type has been modeled. Changes in PPI coupling therefore cannot be confounded by changes in overall signal during both conditions. Moreover, to ensure that changes in noise/signal variability during the two affective contexts cannot confound the PPI analyses, we also compared the signal-to-noise (SNR) ratio within the VMPFC, VS, and insula during threat and safe blocks. To this end, we extracted the filtered signal timeseries from a 12 mm sphere centered on our activation peaks within the VMPFC and insula and from a 6 mm peak centered on our activation peak in VS, for the time periods corresponding to the threat and safe blocks. We then calculated SNR (defined as the absolute condition-specific mean / the condition-specific standard deviation) during safe and threat blocks and compared them statistically. This did not reveal

any SNR differences between safe and threat blocks in VMPFC ($T_{(32)}$ = 0.291, p = 0.773), VS ($T_{(32)}$ = 1.182, p = 0.246), and insula ($T_{(32)}$ = 0.475, p = 0.638), even after controlling for confounding factors, such as age and trait anxiety (VMPFC: affective context coefficient estimate: -0.008, T = -0.30, p = 0.769; VS: affective context coefficient estimate: -0.033, T = -1.20, p = 0.236; insula: affective context coefficient estimate: -0.020, T = -0.48, p = 0.632). This demonstrates that any change in PPI coupling between both affective contexts is not confounded by differences in overall signal or noise levels and therefore most likely reflects changes in functional coupling.

B.4 Results

Threat of shock induces anticipatory anxiety

Analysis of the SCR, behavioral, and neuroimaging data confirmed that our threat-of-shock manipulation was indeed successful in inducing anticipatory anxiety.

First, mean SCRs during decision trials in threat periods were increased relative to safe decision trials. A two-way repeated-measures ANOVA (with Greenhouse-Geisser correction) with factors time (0-16 s following trial onset, averaged across 1-sec bins) and affective context (threat vs. safe) revealed significant main effects for both affective context ($F_{(1,30)}$ = 82.180, p < 0.0001) and time ($F_{(16,30)}$ = 3.245, p < 0.019), as well as a significant interaction of affective context and time ($F_{(1,16)}$ = 31.15, p < 0.001). These SCR increases lasted for the full trial (from 1 until 16 seconds after each trial onset), as indicated by Bonferroni-corrected pairwise t-tests contrasting SCR responses during threat and safe trials (all $T_{(30)}$ = 9.07, all p < 0.0001). Thus, the ongoing threat of painful shocks had a physiologically arousing impact on our participants during performance of the choice paradigm.

Second, SCR increases for threat contexts were also evident during time periods when electrical stimulation was actually administered. ANOVA (with Greenhouse-Geisser correction) of the SCR data following stimulus administration also revealed highly significant main effects of affective context $F_{(1,30)}$ = 48.2, p < 0.0001) and time ($F_{(16,30)}$ = 15.826, p < 0.0001), as well as a significant interaction of affective context and time ($F_{(1,16)}$ = 16.3, p < 0.0001). These SCR increases for strong relative to weak stimulation were evident from 3 until 16 seconds after the onset of the electric stimulus (all $T_{(30)}$ = 6.94, all p < 0.001). This confirms that the strong stimulation itself – on top of the expectation

of this stimulation throughout the block – led to strong physiological arousal in our participants (Figure B-2a).

Third, the arousing impact of the anticipatory anxiety during threat contexts was also confirmed by the choice reaction time (RT) data (Figure B-2b). A regression of RTs on the independent variables affective context (threat versus safe) and choice (LO and SO) revealed significant effects for both affective context ($T_{(32)} = -3.01$, $p = 0.005$) and choice ($T_{(32)} = -4.24$, $p < 0.001$), but no interaction ($T_{(32)} = -0.11$, $p = 0.91$). Participants were faster to respond when choosing LO than SO, and importantly also for choices in the threat versus safe context. The latter result demonstrates the effectiveness of our emotional manipulation on overt behavior (Robinson et al., 2013b).

Finally, the affective efficiency of the painful electric stimulation during threat contexts was also confirmed by the fMRI data. When contrasting neural responses to the strong stimulation during threat blocks with those to the weak stimulation in safe blocks, we observed activation of a putative pain matrix (Tracey and Mantyh, 2007; Leknes and Tracey, 2008) including bilateral insula, bilateral supplementary motor area, cerebellum, and right anterior/mid-cingulate cortex ($p < 0.05$, whole-brain FWE-corrected; Figure B-2c).

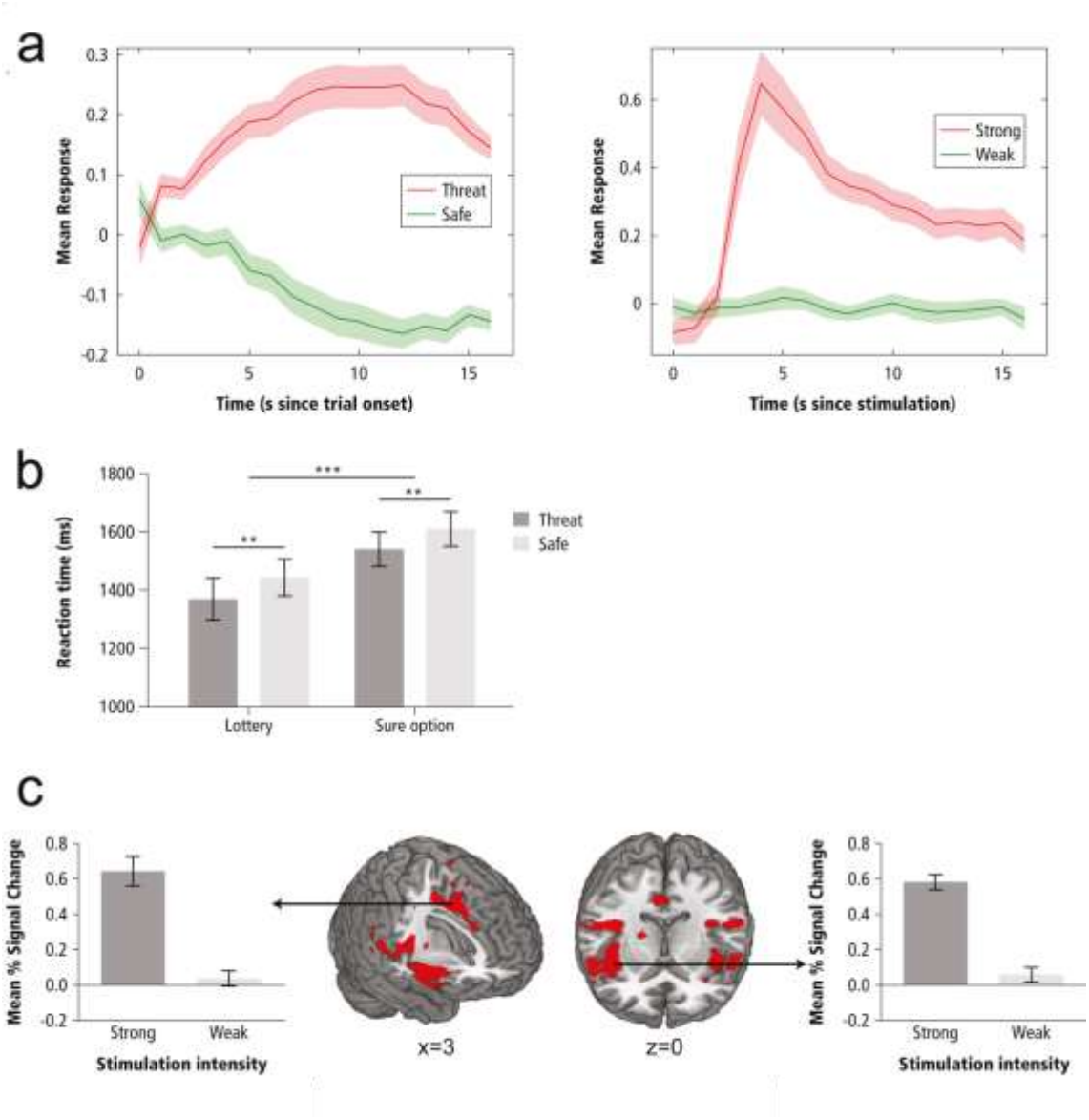

Figure B-2: Evidence for efficacy of the threat-of-shock paradigm in inducing anticipatory anxiety. (a): SCR changes due to threat-of-shock manipulation. The plots show mean (+/- SEM.) SCRs during choices under threat-of-shock (left panel) and during the experience of the electrical shocks (right panel). Significantly increased SCRs were observed for choices during anticipation of strong compared to weak stimulation, and following strong versus weak stimulation. This demonstrates the psychological and physiological efficacy of the threat-of-shock manipulation. (b): Influence of choice and threat-of-shock manipulation on reaction times. RTs were decreased under anticipatory anxiety and for choices of the lottery option. (c): Neural correlates of strong versus weak electrical stimulation. Strong compared to weak stimulation intensity led to activation of the "pain matrix" including bilateral insula, anterior and mid cingulate cortices, right somatosensory cortex and cerebellum ($p < 0.05$, FWE-corrected). Mean (+/- SEM) beta estimates extracted from middle cingulate cortex ($x = 14$, $y = -33$, $z = 46$, left panel) and from right insula ($x = 41$, $y = -1$, $z = -5$) showed signal increase during strong electrical stimulation and no signal increase during weak stimulation. These activation patterns due to strong versus weak electrical stimulation demonstrate the physiological effectiveness of the threat-of-shock manipulation. *** $p < 0.001$, ** $p < 0.01$

Anticipatory anxiety does not affect revealed preferences

The analysis of the choice data showed that anticipatory anxiety did not change overt preferences, as measured with different indices. First, gambling rates did not differ significantly between threat and safe trials ($T_{(32)}$ = 0.05, p = 0.96; Figure B-3a). Subjects were in fact highly consistent in their decisions between threat and safe trials: When specific gambles were offered in both contexts, the same option was chosen in 89% of cases (Figure B-3b). Second, when preferences were modeled via Cumulative Prospect Theory (CPT), we found mild degrees of loss aversion that did not differ between the threat and safe contexts (Figure B-3c), as confirmed via t-test of the difference between λ_S and λ_T (λ_S [SEM] = 1.291 [0.05] and λ_T [SEM] = 1.29 [0.05]; mean (λ_S - λ_T) [SEM] = -0.0007 [0.0085]), ($T_{(32)}$ = 0.08, p = 0.93). Finally, we tested whether the curvature for gains and losses was changed during threat of shocks compared to the safe contexts. We did not observe any significant differences between both contexts for alpha (α_T [SEM]: 0.933 [0.0118], α_S [SEM]: 0.934 [0.0126], $T_{(32)}$ = -0.5052, p = 0.617) and beta (β_T [SEM]: 1.077 [0.0144], β_S [SEM]: 1.078 [0.0134], $T_{(32)}$ = 0.2189, p = 0.8281). Across participants, the small changes in loss aversion and risk attitude due to affective context appeared to be normally distributed (as confirmed by Lilliefors test for normality (Lilliefors, 1967) with λ: D = 0.14, p = 0.13; Figure B-3d; α: D = 0.12, p = 0.29; and β: D = 0.12, p = 0.27). Thus, any changes in loss aversion and risk attitude due to anticipatory anxiety were not significantly different from zero and normally distributed across participants; it is therefore unlikely that different participants reacted in a different manner to the affective manipulation.

The similar choice behavior in the two affective contexts raises the question whether participants may have responded from memory when each lottery was presented for the second time in a different affective context. Such memory effects are generally unlikely, as participants would have to remember their responses across 280 trials, and as the presentation order across affective contexts was fully balanced. Nevertheless, we explicitly tested whether participants pondered their choices in a similar fashion for first and second presentations of lotteries within each context, by means of regression models that included the order of presentation of identical gambles (first vs. second presentation), affective context (threat vs. safe), the experimental trial (1 to 280), gain and loss amounts, as well as relevant two-way interaction terms. Importantly, RTs during first vs second presentations were not significantly different (Order coefficient: -14.438, T = -0.13, p = 0.895) and were similarly affected by affective context (Affective Context * Order coefficient: 40.4184,

T = 1.41, p = 0.159), possible gains (Gains * Order coefficient: -1.809, T = -0.70, p = 0.481) and possible losses (Losses * Order coefficient: -0.148, T= -0.04, p = 0.966). These results suggest that participants did not respond from memory for the second presentations of the lotteries but rather actively evaluated all lotteries. This conclusion is further backed up by our finding of robust value signals in the BOLD data (see below).

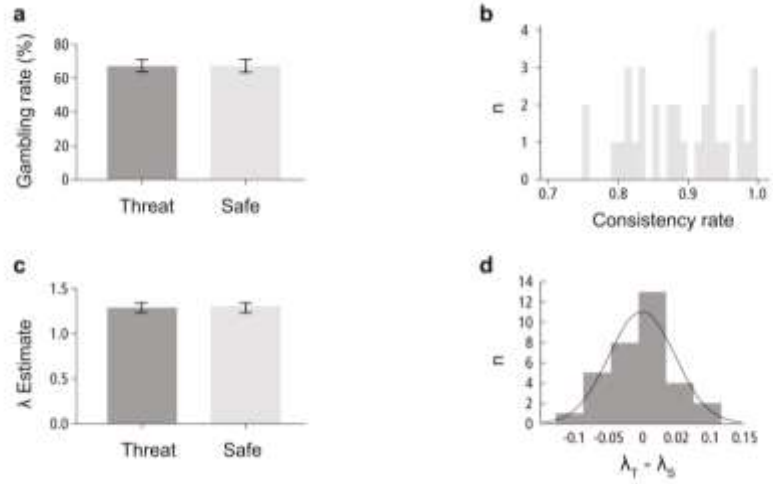

Figure B-3: Effects of the threat-of-shock manipulation on choices. (a): Gambling rates (mean +/- SEM) did not differ for the two affective contexts. (b): Histogram of participants' consistency rates (match of responses [in %] for the identical choices across the affective contexts). Participants chose the same options on average in 89% (+/- 11% SEM) of all gambles that were repeated across contexts. (c): Average (+/- SEM) loss aversion parameter λ across both contexts. Both threat and safe trials led to similar degrees of loss aversion. (d): Histogram of the participants' changes in loss aversion due to the affective manipulation. The changes were not different from zero and normally distributed (mean (λ_S - λ_T) [SEM] = -0.0007 [0.0085]). Thus, the analysis of different behavioral parameters revealed the stability of choice across affective contexts.

Anticipatory anxiety shifts the focus of neural valuation from positive ESV to negative ESV coding

Even though anticipatory anxiety did not change overt behavior, it may nevertheless have affected the way the brain valued the choice options during decision-making. Initial analysis confirmed our expectation that, when pooling across threat and safe trials, ESV correlated positively with activity in the VMPFC (x = 2, y = 47, z = -8, T = 3.34) and the VS (sub-region 1: x_1 = 3, y_1= 11, z_1 = -8, T = 3.59; sub-region 2: x_2 = -5, y_2 = 12, z_2 = -5, T = 3.42). However, when directly comparing neural correlates of ESV between threat and safe trials, we found that anticipatory anxiety led to a significant disruption of ESV coding in both VMPFC (x= -3, y = 39, z = -5, T = 3.59; Figure B-4a and c) and VS (x = 6, y = 6, z = -6; T = 3.24; Figure B-4a and d). To characterize this interaction effect, we investigated parametric correlations between BOLD and ESV during threat and safe trials separately.

Both the VMPFC and VS coded ESV during safe trials (VMPFC: x = -3, y = 40, z = -8, T = 3.87; VS: x = 6, y = 6, z = -6, T = 4.56) but not during threat trials (no significant ESV-related activity was found in those regions). Moreover, a conjunction analysis of the ESV contrast images during threat and safe trials did not reveal any activity in these regions.

A very different pattern of results was observed in right anterior insula, where ESV coding also differed between both affective contexts (x = 39, y = 0, z = 0; T = 3.25; Figure B-4b and e). Separate analyses of neural signals during both contexts revealed that ESV was not significantly correlated with BOLD signals during safe trials (no voxel was found), but was instead negatively correlated with BOLD signals during threat trials (left insula cluster 1: x = -30, y = 21, z = 13**,** T = 3.48; left insula cluster 2: x = -40, y = 20, z = 4**,** T = 3.38; right insula: x = 39, y = -12, z = -6, T = 3.44). Again, a conjunction analysis of the ESV contrast images during threat and safe trials did not reveal any activity in this area. Thus, during anticipatory anxiety, the insula increased its activity for smaller U(x) and therefore larger potentially negative outcomes. Anticipatory anxiety therefore led to a qualitatively distinct style of neural value coding, by disrupting positive value coding in the VMPFC and the VS while simultaneously enhancing negative value coding in the anterior insula.

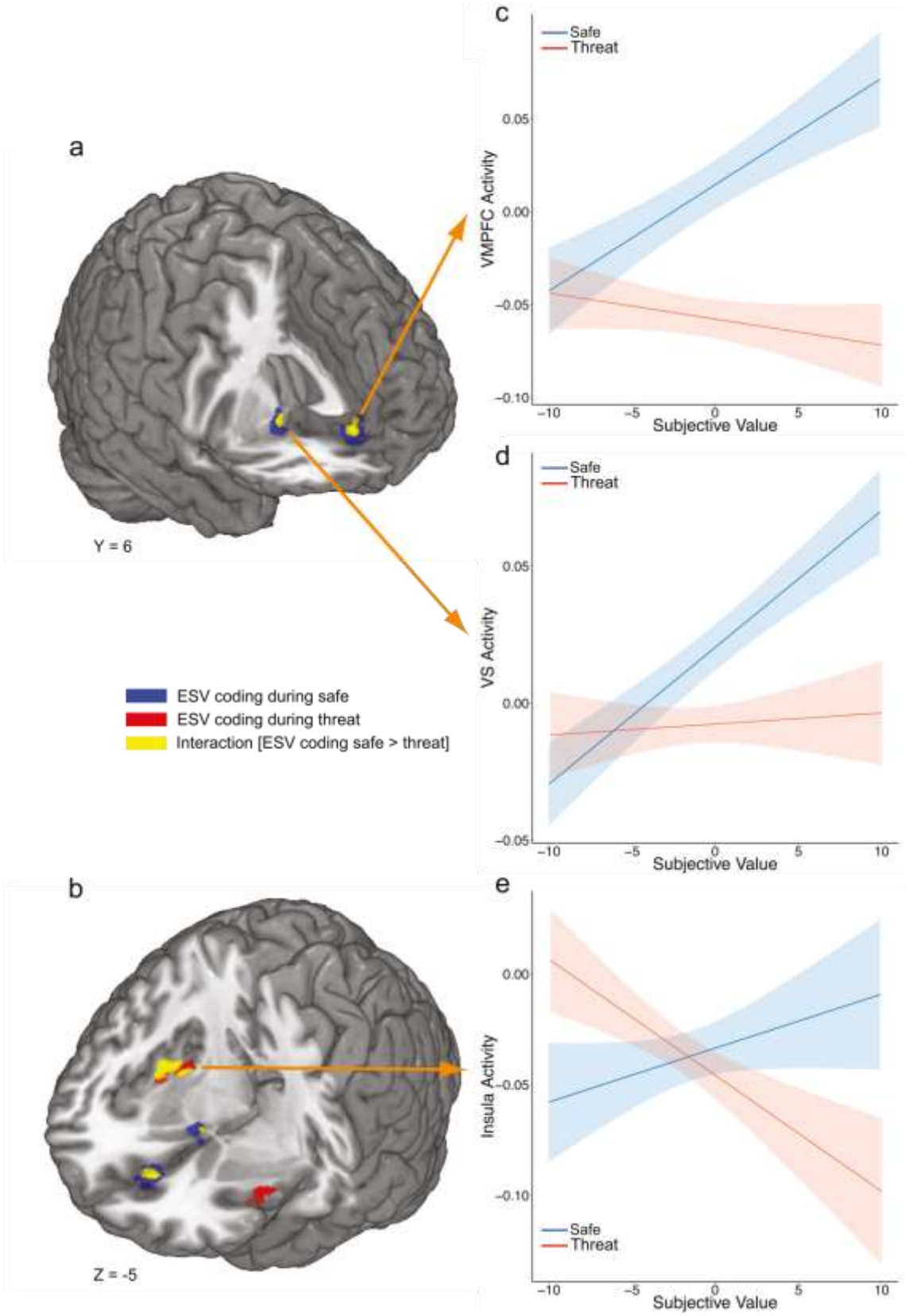

Figure B-4: Changes in the correlation of regional BOLD signals with subjective values due to threat-of-shock (interaction contrast $ESV_S > ESV_T$). Such changes were observed in (a) VMPFC (x = -3, y = 39, z = -5), and VS (x = 6, y = 6, z = -6) as well as (b) insula (x = 39, y = 0, z = 0). Voxels in blue show positive correlation with subjective value during safe trials, voxels in red show negative correlations with

SV during threat trials, and voxels in yellow show a significant interaction. (c-e) ROI analyses confirm the pattern of interactions found in the SPM analyses. To avoid circularity, a LOSO approach was employed to extract beta weights from independent ROIs (see text for details). The plots show the average trial-by-trial relationship (lines +/- SEM as shaded area) between regional activity and subjective value, as estimated in regression analyses of these betas (see main text). For both VMPFC (c) and VS (d), positive correlations between subjective value and BOLD activity were present in safe trials, but were significantly weaker (and in fact absent) during threat trials. In contrast, the insula (e) showed no correlation with subjective value during safe trials, but significantly stronger negative value coding (increasing activity for decreasing expected subjective value) during threat trials. These findings suggest that anticipatory anxiety caused a distinct style of neural value coding, by disrupting positive value coding in the VMPFC and the VS and simultaneously enhancing negative value coding in the anterior insula. All imaging results are displayed at $p < 0.005$ for display purposes. The shaded error bars reflect robust and clustered SEM.

Anticipatory anxiety changes trial-by-trial relationship between neural signal in ROIs and economic choice

The fMRI results described above identified regions in the VMPFC, VS, and insula where neural activity correlated with different aspects of subjective values during the two affective contexts (positive ESVs during safe trials and negative ESVs during threat trials). As a next step, we corroborated that the change in neural coding in these regions between both contexts is tightly linked to behavior. For this analysis, we extracted the neural activity for each trial from the ROIs in the VMPFC, VS, and insula and regressed the choice outcomes for each trial on the activity in each region, the affective context, and the interaction of both variables (see section "Methods and materials"). Prior to this analysis, we ensured that the ROI data could replicate the results of the SPM analysis, by performing a regression of neural activity in each region on subjective values, affective context, and their interaction. Importantly, the ROIs for both analyses were similarly defined using the well-established leave-one-subject-out approach (Esterman et al., 2010), which avoids circularity and enabled us to perform this prediction analysis with independent data (see section "Methods and materials").

The independent ROI analysis clearly confirmed the results of the SPM analysis. As illustrated in Figure B-4, all regions showed a negative change in value coding from the safe to the threat context, as indicated by a significant interaction between subjective value and threat [VMPFC: interaction coefficient = -0.0071, T = -2.70, p = 0.0035; VS: interaction coefficient = -0.0046, T = -1.99, p = 0.0231; insula: interaction coefficient = -0.0077, T = -2.01, p = 0.0221]. Importantly, the pattern of this interaction was different for the insula compared to the ventral striatum and the VMPFC (Figure B-4): Whereas the signal in VMPFC and VS tracked ESVs during the safe condition and was suppressed under conditions of threat (VMPFC: ESV_S coefficient = 0.0057, T = 2.77, p = 0.0028;

ESV_T coefficient = -0.0014, T = -0.77, p = 0.2201; VS: ESV_S coefficient = 0.0050, T = 3.95, p < 0.0001, ESV_T coefficient = 0.0004, T = 0.25, p = 0.4028), activity in the insula did not track value during safe trials but coded negative subjective value during the threat condition (insula: ESV_S coefficient = 0.0024, T = 0.86, p = 0.194; ESV_T coefficient = -0.0052, T = -1.98, p = 0.0239]. These results from the independent ROIs thus fully confirm the relationship between subjective values and regional signal in VMPFC, VS and insula.

Next, we tested whether trial-by-trial signal changes within our ROIs in the valuation network directly map onto actual economic decisions. We predicted that the trial-by-trial signal within these independent regions not only reflects subjective values, but can also directly predict observed choices. In line with all previous results, we expected that decision outcomes in the absence of threat should be predicted by only the VMPFC and VS signal, but in the presence of threat only by insula activity. In line with this prediction, we found significant interactions between neural signal and threat for VMPFC (interaction coefficient = -0.246, Wald Z = -2.54, p = 0.0056) and insula (interaction coefficient = -0.1162, Wald Z = -1.71, p = 0.044). As illustrated in Figure B-5a, VMPFC neural signals predicted choice only during safe trials (safe coefficient = 0.1876, Wald Z = 3.35, p = 0.0004) but not during threat trials (threat coefficient = -0.0550, Wald Z = -0.79, p = 0.2161), whereas the opposite pattern was obtained in the insula (safe coefficient = 0.0302, Wald Z = 0.69, p = 0.24475; threat coefficient = -0.089, Wald Z = -1.84; p = 0.033). The corresponding interaction in the ventral striatum was weaker and non-significant (interaction coefficient = -0.125, Wald Z = -0.82, p = 0.206), perhaps due to greater noise in the signals extracted from this small region. Nevertheless, in analogy to the VMPFC and again, consistent with our hypothesis, neural signals in the VS predicted choices during safe trials (safe coefficient = 0.2189, Wald Z = 2.60, p = 0.0047) but not during threat trials (threat coefficient = 0.0909, Wald Z = 0.95, p = 0.17095, Figure B-5a). Taken together, these regression analyses therefore establish that the different patterns of neural ESV coding in the two affective contexts are both clearly linked to choice behavior. Moreover, all our results on neural valuation signals show that participants did compute subjective values for all choices, even though these computations were implemented in different cortical areas depending on affective context. The presence of these context-dependent value computations makes it very unlikely that participants responded from memory on the 50% of trials where gambles were presented a second time.

Negative baseline shifts in VMPFC during anticipatory anxiety

What mechanism may underlie the breakdown of positive value-coding under anticipatory anxiety? One possibility is that in the threat context, processing of lottery values in the VS and VMPFC may interact with the negative emotional value associated with the anticipated uncomfortable stimulation (Talmi et al., 2009; Winecoff et al., 2013). Given the salience of anticipatory anxiety relative to lottery outcomes, the processing of the aversive emotional event can be expected to overshadow value coding with respect to probabilistic financial gains. If this were the case, then BOLD signals in the VS and VMPFC should be lower on average for threat compared to safe trials, pooled across all possible financial outcomes. We tested this conjecture by contrasting task-related activity in threat versus safe trials, independent of correlations with trial-wise ESV. Consistent with our conjecture, threat trials were associated with decreased task-related activations in VMPFC (x = -8, y = 35, z = -2; T = 5.58; Figure B-5b), amygdala (x = 29, y = 3, z = -29; T = 4.56) and insula (x = 45, y = 0, z = 1; T = 3.64, x = -45, y = -3, z = -2; T = 4.45). These observations confirm that anticipatory anxiety due to the impending strong stimulation can decrease the overall activation of valuation brain regions (Talmi et al., 2009). This may reduce the excitability of these brain regions for coding the positive values of probabilistic financial gains, as suggested by the observed breakdown of positive ESV coding in the threatening context.

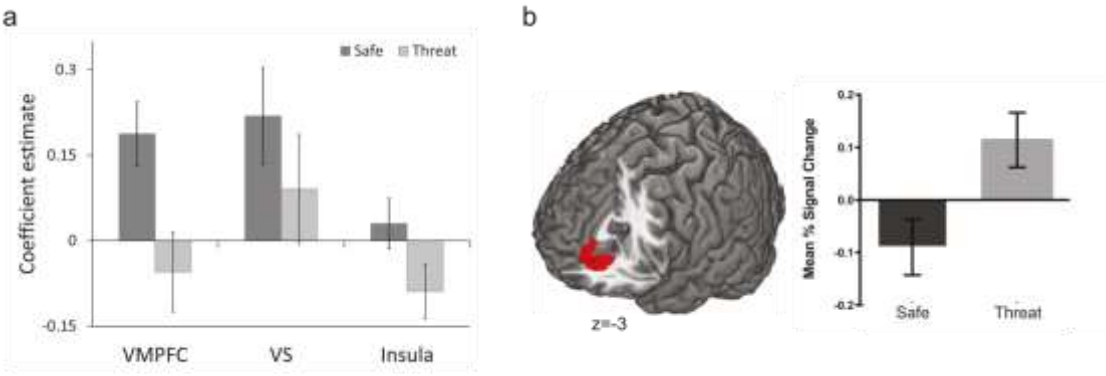

Figure B-5: (a) Results from logistic regressions predicting choice based on regional activity in VMPFC, VS, and insula. Both VMPFC and VS predict choice only during safe trials, but not during threat trials. In contrast, the insula shows the opposite pattern of predicting choice (negatively) only during threat trials, but not during safe trials. (b) Decreases in average VMPFC BOLD signals during choices due to threat-of-shock ($T_{Safe} > T_{Threat}$). Anticipation of strong electrical stimulation led to general BOLD signal reduction in VMPFC (x = -8, y = 35, z = -2) pooled across all ESV levels. Anticipatory anxiety led to decreases in overall activity for valuation brain regions during choice, possibly due to the inherent negative value of the impending strong stimulation. The bar plot shows mean regression coefficients extracted from a 12 mm sphere around the VMPFC activation peak to illustrate the pattern of suppressed activity during choices in the threat vs safe context. The imaging results are displayed at $p < 0.005$ for display purposes. Error bars reflect SEM.

Anticipatory anxiety decreases functional connectivity in the valuation network

If anticipatory anxiety disrupts the neural processing of possible financial outcomes during choice, then this may not only affect activity in single brain areas but also functional communication within the connected network of value-related brain areas. Thus, we hypothesized that the VMPFC may show reduced functional connectivity with other value-related brain regions during anticipatory anxiety. We formally tested this conjecture with psychophysiological interaction (PPI) analyses of context-dependent changes in functional connectivity of the VMPFC (seeded at x = -3, y = 39, z = -5, Figure B-6a) between threat and safe contexts. This revealed that threat contexts indeed lead to significantly decreased functional coupling of the VMPFC with two other key value-related brain areas: the VS (x = -6, y = 14, z = -11, T = 3.64, Figure B-6b and c) and the insula (x = -38, y = 17, z = -8, T = 4.9; Figure B-6b and d). This confirms that anticipatory anxiety disrupted positive ESV coding both at the level of local activity within areas of the value network, but also at the level of functional integration between these brain regions.

We also tested the hypothesis that anticipatory anxiety may change functional communication between valuation areas and regions outside the classic valuation system, which may possibly be involved in action selection irrespective of affective context. For this purpose, we conducted PPI analyses based on seed regions in VS (at x = 6, y = 6, z = -6), the insula (at x = 39, y = 0, z = 0), and the dorsolateral PFC (seeded at the DLPFC peak voxel in the neurosynth automated meta-analysis for the term cognitive control: x = 50, y = 36, z = 24). However, we did not find any reduction in functional connectivity between our areas of interest and other regions outside the classic valuation system. The unchanged functional involvement of cognitive control areas such as the DLPFC may perhaps reflect that neurons in the VMPFC and insula not only compute values but also encode signals used for flexible choice and action selection (Paulus et al., 2005; Hunt et al., 2014; Strait et al., 2014). In any case, our results confirms that anticipatory anxiety during the threatening context disrupted positive ESV coding both at the level of local activity within areas of the value network, but also at the level of functional integration between these brain regions.

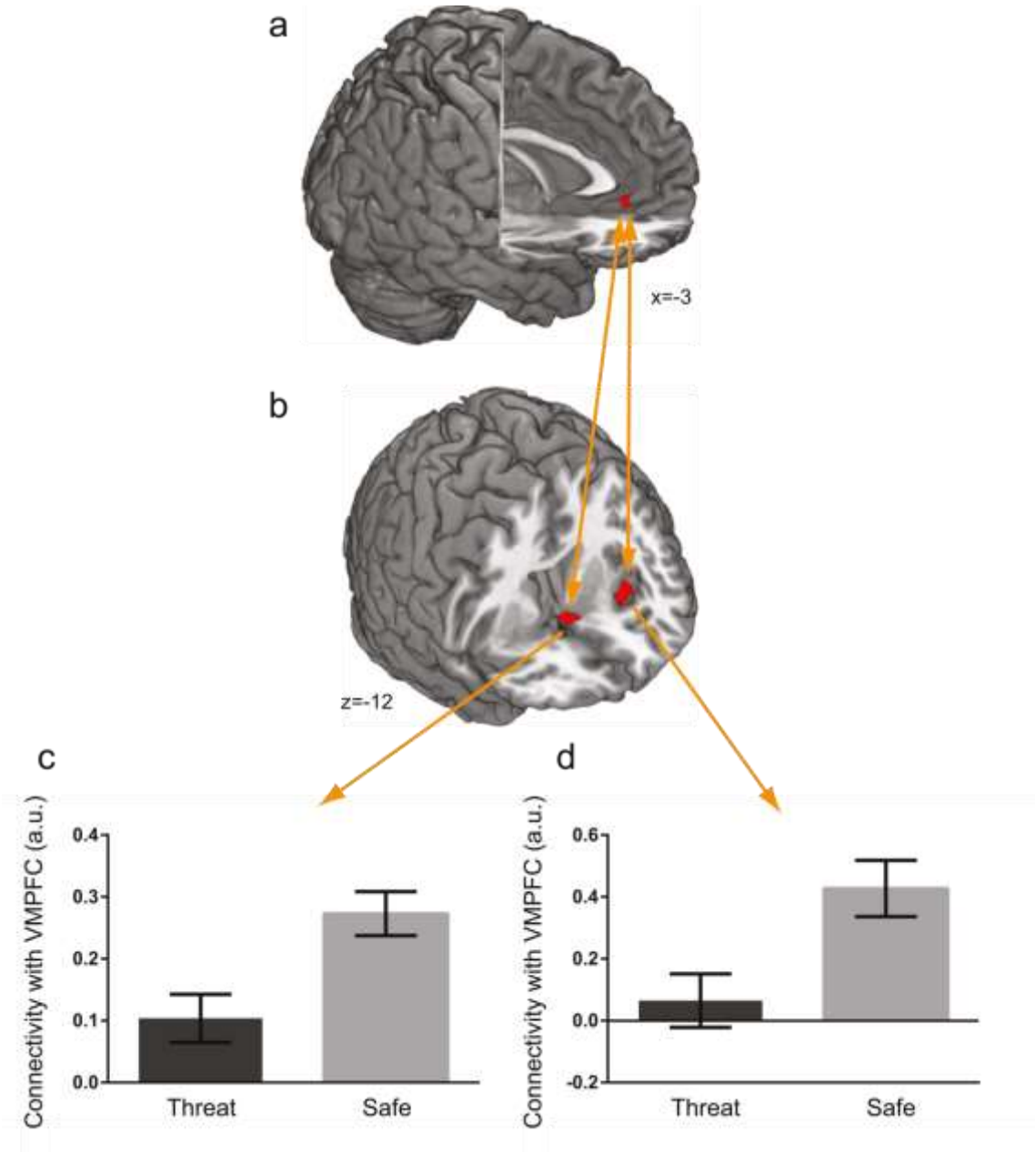

Figure B-6: Changes in functional connectivity within the valuation network during anticipatory anxiety. Threat-of-shock led to decreased connectivity between (a) the seed region in VMPFC (x = -3, y = 39, y = -5); and both VS (x = -6, y = 14, z = -11, T = 3.64; Figure B-6b and c) and insula (x = -38, y = 17, z = -8, T = 4.9; Figure B-6b and d). Thus, anticipatory anxiety disrupted the functional integration of the brain's valuation network. Error bars reflect SEM. All results are displayed at $p < 0.005$ for display purposes. Error bars reflect SEM.

B.5 Discussion

This study investigated how incidental anxiety can affect value-based decisions under risk. Purely behavioral effects of background emotions on choice are well-documented in the clinical and psychology literature (Maner et al., 2007; Clark et al., 2012; Giorgetta et al., 2012; Grecucci et al., 2013) but have been less prominently investigated from the perspective of underlying neural activity (Pessoa and Pereira, 2013). Here we investigated this issue with an experimental threat-of-shock paradigm that allowed us to induce anticipatory anxiety in healthy participants while they took risky choices during fMRI. We find that this affective manipulation did not change overt choices, but fundamentally altered value-coding in the brain: In the safe context, the expected subjective values of the gambles were coded positively in the VMPFC and VS. During the threat context, these processes collapsed and were replaced by negative value coding in the anterior insula (i.e., increased activity for increasingly worse outcomes). In general, our results are congruent with previous findings (Kim et al., 2011) that the VS and VMPFC can code subjective value in a positive manner, whereas the insula can show negative correlations with subjective values. However, our data show that these two types of value representations can be dissociated and may depend flexibly on the affective state, with a dominance of positive value coding in the VMPFC and VS during neutral states, and a prevalence of negative value coding in the insula during negative affective states. This dependence of value coding on affective context may not have been evident in previous neuroimaging studies that measured value-related brain activity in just one affective context (usually neutral; for review, see Levy and Glimcher, 2011, 2012; Bartra et al., 2013). However, please note that both the diminished value sensitivity in VMPFC and its decreased communication with the striatum under anticipatory anxiety are clearly consistent with rodent studies that have demonstrated impaired value coding and reward learning under stress (Dias-Ferreira et al., 2009).

Strikingly, the shift in value-coding between affective contexts was not accompanied by changes in revealed economic preferences, as measured by overt choice (only reaction times were affected). Nevertheless, our analyses revealed a tight link between neural value signals and behavior in both affective contexts, with neural activity in VMPFC and VS predicting choices only for safe trials and neural activity in the insula predicting choices only for threat trials. Our results thus demonstrate that the same behavioral outcome can be realized through different neural value-coding mechanisms. This finding is generally

congruent with multiple-systems theories of decision-making stating that different valuation systems can guide choice depending on the properties of the decision-making situation (Rangel et al., 2008). That neural valuation processes can change during negative affective states is corroborated by findings (Schwabe et al., 2012) that hormonally induced stress can lead to a change from goal-directed to habitual decision control and reduced value sensitivity in medial prefrontal cortex. Our findings therefore suggest that during anxiety, the same behavioral choices may have been brought about by goal-directed negative value computations in the insula (rather than positive value computations in the VMPC and VS), possibly in conjunction with compensatory mechanisms such as heuristics and increased habitual responding. In either case, we speculate that the reduction of value coding in the VMPFC and VS during anxiety may indicate a decrease in the positive value of the available choice options, even though such speculations need to be taken with caution given the problem of reverse inference in neuroimaging studies (Poldrack, 2011).

The conjecture that anticipatory anxiety may have decreased neural sensitivity for the subjective value of the choice-relevant stimuli is congruent with our findings of (1) a general negative baseline shift in VMPFC during threat and (2) reduced functional connectivity of this area with the insula and the VS. This assumption dovetails with results of previous neuroimaging studies on the integration of emotional and economic value (Talmi et al., 2009; Park et al., 2011; Winecoff et al., 2013). Specifically, recent findings suggest that VMPFC signals track the emotional value of stimuli on the one hand (Winecoff et al., 2013) and economic value on the other (Levy and Glimcher, 2012). Moreover, Talmi et al. (2009) found decreased value signals in VMPFC for choices comprising a task-contingent painful stimulus. Here we show that even a fully incidental but highly salient and emotionally aversive event – the anticipation of painful electrical stimulation – can overshadow neural signals related to economic value computations. This supports the notion that emotional and economic value may be processed by the same VMPFC region, extending the hypothesis that the VMPFC encodes a standardized economic value signal to the domain of emotional value (Montague and Berns, 2002; Rangel and Hare, 2010). Moreover, our findings are consistent with suggestions that the VMPFC may compute subjective values even for certain task-irrelevant stimuli (Lebreton et al., 2009; Levy et al., 2011; Kühn and Gallinat, 2012) and that negative affective states (i.e., anxiety or stress) can lead to structural and functional changes in the VMPFC (Liston et al., 2006; McEwen, 2007).

However, it could be argued that the diminished VMPFC activity under anticipatory anxiety may not reflect changes in neural value processing but rather reduced default-mode network (DMN) activity, as DMN activity is commonly observed in the VMPFC. However, this alternative explanation appears very unlikely: Inhibitions of DMN-activity would be expressed in a whole network of regions comprising the VMPFC, PCC, and TPJ (Buckner et al., 2008; Andrews-Hanna et al., 2010; Li et al., 2014), would be accompanied by a concomitant increase in fronto-parietal activity (Fox et al., 2005; Sridharan et al., 2008; Chen et al., 2013), and would be more marked in conditions with higher cognitive demand and thus longer reaction times (McKiernan et al., 2003). In our task, a negative baseline shift was found only in the VMPFC (but not the PCC and TPJ), in the threat context that actually led to faster RTs (rather than increased demand indicated by slowed responses). The reductions in VMPFC signals found here are therefore more likely to indicate reduced positive value coding due to the high negative value of the impending painful stimulation.

Our results show that the functional changes due to the threat of shock not only affected the VMPFC in isolation but also its functional connectivity with other areas. This suggests weakened function of an entire valuation network (rather than single regions) under anticipatory anxiety. Viewed from this network perspective, the reduction in connectivity between VMPFC and insula may indicate a release from inhibition in the insula that may have resulted in stronger negative value-coding during threat trials. Importantly, the insula showed two distinct types of signals in our study. One of these signals appeared related to valuation during threat trials and consisted of a negative baseline shift and enhanced negative value coding (Sescousse et al., 2013; Tanaka et al., 2014). The other signal was elicited by the painful shocks when these were actually administered, reminiscent of the well-reported activity in the so-called pain matrix (Tracey and Mantyh, 2007; Leknes and Tracey, 2008). We speculate that these two effects may reflect the involvement of separate neural populations, for two reasons. First, it is well established that the insula contains functionally diverse types of neurons (Chang et al., 2013; Uddin et al., 2014) that are nevertheless spatially intermingled. This may explain why in our data, the activation peaks for both types of signals were spatially separated (5 mm apart). Second, pain-processing neurons in the insula usually show positive baseline shifts during the anticipation and actual delivery of (predictable) painful shocks (Jones et al., 1992; Apkarian et al., 2005; Dalton et al., 2005; Carlsson et al., 2006; Choi et al., 2012). The negative baseline shifts found here during the threat blocks (with unpredictable stimulation) are thus more likely to reflect valuation processes, in line with findings that insula activity reflects

value computations in other contexts (Sescousse et al., 2013; Tanaka et al., 2014) and is susceptible to stress during decision-making (Lighthall et al., 2009, 2012).

Our study may have important implications for neurobiological perspectives on maladaptive anxiety and related psychopathology. Transient anxiety states normally carry adaptive value since they may increase vigilance and attention to possible negative outcomes that should be avoided. This functional anxiety, however, may turn into a maladaptive state if anxious behavior is permanently adopted and becomes detached from the environment (Grupe and Nitschke, 2013). We find that transient anxiety leads to a change in neural value coding in the absence of changes in overt behavior. This situation may resemble prodromal stages of anxiety disorders where no pathological decision-making is evident, as the anxious persons might use compensatory strategies (e.g., increased focus on task or calculation) in order to overcome the pathological risk-aversion typical for anxiety disorders (Maner et al., 2007; Giorgetta et al., 2012). Our findings suggest that during such early stages, a neural focus on negative value might not be accompanied by striking behavioral changes. However, the permanent adoption of these strategies might lead to gradual deterioration of the neural valuation system and/or compensatory strategies until behavioral deviations emerge. Similar mechanisms may be at play in depression, where early stages of the illness are associated with decreased hedonic capacity (indicated by changes in neural value processing) without observable aberrant decision-making, whereas later stages are also often associated with altered value-based decision making (Treadway and Zald, 2011; Der-Avakian and Markou, 2012). This may be consistent with the assumption that a persistent focus on negative value coding due to the ongoing incidental negative affect may ultimately lead to a functional deterioration of the brain's valuation system and therefore changes of overt behavior. Thus, our findings suggest a pathway for the transition from adaptive to maladaptive emotional responses, as commonly observed in anxiety and affective disorders (Paulus and Yu, 2012).

B.6 References

Andersson JL, Hutton C, Ashburner J, Turner R, Friston K (2001) Modeling geometric deformations in EPI time series. NeuroImage 13:903–919.

Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL (2010) Functional-anatomic fractionation of the brain's default network. Neuron 65:550–562.

Apkarian AV, Bushnell MC, Treede R-D, Zubieta J-K (2005) Human brain mechanisms of pain perception and regulation in health and disease. Eur J Pain Lond Engl 9:463–484.

Bach DR, Flandin G, Friston KJ, Dolan RJ (2009) Time-series analysis for rapid event-related skin conductance responses. J Neurosci Methods 184:224–234.

Bartra O, McGuire JT, Kable JW (2013) The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage 76:412–427.

Berns GS, Capra CM, Chappelow J, Moore S, Noussair C (2008) Nonlinear neurobiological probability weighting functions for aversive outcomes. NeuroImage 39:2047–2057.

Borkenau P, Ostendorf F (1993) NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa und McCrae. Göttingen: Hogrefe.

Bruhin A, Fehr-Duda H, Epper T (2010) Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. Econometrica 78:1375–1412.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1–38.

Capra CM, Jiang B, Engelmann JB, Berns GS (2013) Can Personality Type Explain Heterogeneity in Probability Distortions? J Neurosci Psychol Econ 6:151–166.

Carlsson K, Andersson J, Petrovic P, Petersson KM, Ohman A, Ingvar M (2006) Predictability modulates the affective and sensory-discriminative neural processing of pain. NeuroImage 32:1804–1814.

Chang LJ, Yarkoni T, Khaw MW, Sanfey AG (2013) Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. Cereb Cortex N Y N 1991 23:739–749.

Chen AC, Oathes DJ, Chang C, Bradley T, Zhou Z-W, Williams LM, Glover GH, Deisseroth K, Etkin A (2013) Causal interactions between fronto-parietal central executive and default-mode networks in humans. Proc Natl Acad Sci 110:19944–19949.

Choi JM, Padmala S, Pessoa L (2012) Impact of state anxiety on the interaction between threat monitoring and cognition. NeuroImage 59:1912–1923.

Clark L, Li R, Wright CM, Rome F, Fairchild G, Dunn BD, Aitken MRF (2012) Risk-avoidant decision making increased by threat of electric shock. Psychophysiology 49:1436–1443.

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res Int J 29:162–173.

Dalton KM, Kalin NH, Grist TM, Davidson RJ (2005) Neural-cardiac coupling in threat-evoked anxiety. J Cogn Neurosci 17:969–980.

Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. NeuroImage 19:430–441.

De Martino B, Kumaran D, Seymour B, Dolan RJ (2006) Frames, biases, and rational decision-making in the human brain. Science 313:684–687.

Der-Avakian A, Markou A (2012) The neurobiology of anhedonia and other reward-related deficits. Trends Neurosci 35:68–77.

Dias-Ferreira E, Sousa JC, Melo I, Morgado P, Mesquita AR, Cerqueira JJ, Costa RM, Sousa N (2009) Chronic Stress Causes Frontostriatal Reorganization and Affects Decision-Making. Science 325:621–625.

Di Martino A, Scheres A, Margulies DS, Kelly AMC, Uddin LQ, Shehzad Z, Biswal B, Walters JR, Castellanos FX, Milham MP (2008) Functional connectivity of human striatum: a resting state FMRI study. Cereb Cortex N Y N 1991 18:2735–2747.

Dreher J-C, Kohn P, Kolachana B, Weinberger DR, Berman KF (2009) Variation in dopamine genes influences responsivity of the human reward system. Proc Natl Acad Sci 106:617–622.

Engelmann JB, Maciuba B, Vaughan C, Paulus MP, Dunlop BW (2013) Posttraumatic stress disorder increases sensitivity to long term losses among patients with major depressive disorder. PloS One 8:e78292.

Esterman M, Tamber-Rosenau BJ, Chiu Y-C, Yantis S (2010) Avoiding non-independence in fMRI data analysis: Leave one subject out. NeuroImage 50:572–576.

Fossati P (2012) Neural correlates of emotion processing: from emotional to social brain. Eur Neuropsychopharmacol J Eur Coll Neuropsychopharmacol 22 Suppl 3:S487–S491.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Essen DCV, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A 102:9673–9678.

Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ (1994) Statistical parametric maps in functional imaging: A general linear approach. Hum Brain Mapp 2:189–210.

Giorgetta C, Grecucci A, Zuanon S, Perini L, Balestrieri M, Bonini N, Sanfey AG, Brambilla P (2012) Reduced risk-taking behavior as a trait feature of anxiety. Emotion 12:1373–1383.

Gitelman DR, Penny WD, Ashburner J, Friston KJ (2003) Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. NeuroImage 19:200–207.

Grecucci A, Giorgetta C, Brambilla P, Zuanon S, Perini L, Balestrieri M, Bonini N, Sanfey AG (2013) Anxious ultimatums: how anxiety disorders affect socioeconomic behaviour. Cogn Emot 27:230–244.

Grupe DW, Nitschke JB (2013) Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective. Nat Rev Neurosci 14:488–501.

Harle KM, Allen JJB, Sanfey AG (2010) The Impact of Depression on Social Economic Decision-making. J Abnorm Psychol 119:440–446.

Harlé KM, Chang LJ, van 't Wout M, Sanfey AG (2012) The neural mechanisms of affect infusion in social economic decision-making: a mediating role of the anterior insula. NeuroImage 61:32–40.

Hartley CA, Phelps EA (2012) Anxiety and decision-making. Biol Psychiatry 72:113–118.

Hautzinger M, Bailer M, Worall H, Keller F (1995) Beck-Depressions-Inventar (BDI). Testhandbuch, 2nd ed. Bern: Hans Huber.

Henson RNA, Price CJ, Rugg MD, Turner R, Friston KJ (2002) Detecting Latency Differences in Event-Related BOLD Responses: Application to Words versus Nonwords and Initial versus Repeated Face Presentations. NeuroImage 15:83–97.

Hsu M, Bhatt M, Adolphs R, Tranel D, Camerer CF (2005) Neural systems responding to degrees of uncertainty in human decision-making. Science 310:1680–1683.

Hunt LT, Dolan RJ, Behrens TEJ (2014) Hierarchical competitions subserving multi-attribute choice. Nat Neurosci 17:1613–1622.

Huys QJ, Pizzagalli DA, Bogdan R, Dayan P (2013) Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. Biol Mood Anxiety Disord 3:12.

Jenison RL, Rangel A, Oya H, Kawasaki H, Howard MA (2011) Value Encoding in Single Neurons in the Human Amygdala during Decision Making. J Neurosci 31:331–338.

Jones AK, Friston K, Frackowiak RS (1992) Localization of responses to pain in human cerebral cortex. Science 255:215–216.

Kahneman D, Tversky A (1979) Prospect Theory: An Analysis of Decision under Risk. Econometrica 47:263.

Kang MJ, Rangel A, Camus M, Camerer CF (2011) Hypothetical and real choice differentially activate common valuation areas. J Neurosci Off J Soc Neurosci 31:461–468.

Kelly C, de Zubicaray G, Di Martino A, Copland DA, Reiss PT, Klein DF, Castellanos FX, Milham MP, McMahon K (2009) L-dopa modulates functional connectivity in striatal cognitive and motor networks: a double-blind placebo-controlled study. J Neurosci Off J Soc Neurosci 29:7364–7378.

Kim H, Shimojo S, O'Doherty JP (2011) Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. Cereb Cortex N Y N 1991 21:769–776.

Knutson B, Wimmer GE, Kuhnen CM, Winkielman P (2008) Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. Neuroreport 19:509–513.

Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12:535–540.

Kuhnen CM, Knutson B (2005) The neural basis of financial risk taking. Neuron 47:763–770.

Kühn S, Gallinat J (2012) The neural correlates of subjective pleasantness. NeuroImage 61:289–294.

Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M (2009) An automatic valuation system in the human brain: evidence from functional neuroimaging. Neuron 64:431–439.

Lee D, Seo H, Jung MW (2012) Neural basis of reinforcement learning and decision making. Annu Rev Neurosci35:287-308.

Leknes S, Tracey I (2008) A common neurobiology for pain and pleasure. Nat Rev Neurosci 9:314–320.

Levallois C, Clithero JA, Wouters P, Smidts A, Huettel SA (2012) Translating upwards: linking the neural and social sciences via neuroeconomics. Nat Rev Neurosci 13:789–797.

Levy DJ, Glimcher PW (2011) Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain. J Neurosci 31:14693–14707.

Levy DJ, Glimcher PW (2012) The root of all value: a neural common currency for choice. Curr Opin Neurobiol 22:1027–1038.

Levy I, Lazzaro SC, Rutledge RB, Glimcher PW (2011) Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. J Neurosci Off J Soc Neurosci 31:118–125.

Levy I, Snell J, Nelson AJ, Rustichini A, Glimcher PW (2010) Neural representation of subjective value under risk and ambiguity. J Neurophysiol 103:1036–1047.

Lighthall NR, Mather M, Gorlick MA (2009) Acute Stress Increases Sex Differences in Risk Seeking in the Balloon Analogue Risk Task. PLoS ONE 4:e6002.

Lighthall NR, Sakaki M, Vasunilashorn S, Nga L, Somayajula S, Chen EY, Samii N, Mather M (2012) Gender differences in reward-related decision processing under stress. Soc Cogn Affect Neurosci 7:476–484.

Lilliefors HW (1967) On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. J Am Stat Assoc 62:399.

Lin A, Adolphs R, Rangel A (2012) Social and monetary reward learning engage overlapping neural substrates. Soc Cogn Affect Neurosci 7:274–281.

Liston C, Miller MM, Goldwater DS, Radley JJ, Rocher AB, Hof PR, Morrison JH, McEwen BS (2006) Stress-induced alterations in prefrontal cortical dendritic morphology predict selective impairments in perceptual attentional set-shifting. J Neurosci Off J Soc Neurosci 26:7870–7874.

Litt A, Plassmann H, Shiv B, Rangel A (2011) Dissociating valuation and saliency signals during decision-making. Cereb Cortex N Y N 1991 21:95–102.

Li W, Mai X, Liu C (2014) The default mode network and social understanding of others: what do brain connectivity studies tell us. Front Hum Neurosci 8:74.

Maner JK, Richey JA, Cromer K, Mallott M, Lejuez CW, Joiner TE, Schmidt NB (2007) Dispositional anxiety and risk-avoidant decision-making. Personal Individ Differ 42:665–675.

McEwen BS (2007) Physiology and neurobiology of stress and adaptation: central role of the brain. Physiol Rev 87:873–904.

McKiernan KA, Kaufman JN, Kucera-Thompson J, Binder JR (2003) A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. J Cogn Neurosci 15:394–408.

McLaren DG, Ries ML, Xu G, Johnson SC (2012) A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. NeuroImage 61:1277–1286.

Minassian A, Paulus MP, Perry W (2004) Increased sensitivity to error during decision-making in bipolar disorder patients with acute mania. J Affect Disord 82:203–208.

Minati L, Grisoli M, Franceschetti S, Epifani F, Granvillano A, Medford N, Harrison NA, Piacentini S, Critchley HD (2012a) Neural signatures of economic parameters during decision-making: a functional MRI (FMRI), electroencephalography (EEG) and autonomic monitoring study. Brain Topogr 25:73–96.

Minati L, Grisoli M, Seth AK, Critchley HD (2012b) Decision-making under risk: a graph-based network analysis using functional MRI. NeuroImage 60:2191–2205.

Montague PR, Berns GS (2002) Neural Economics and the Biological Substrates of Valuation. Neuron 36:265–284.

Park SQ, Kahnt T, Rieskamp J, Heekeren HR (2011) Neurobiology of Value Integration: When Value Impacts Valuation. J Neurosci 31:9307–9314.

Patton JH, Stanford MS, Barratt ES (1995) Factor structure of the Barratt impulsiveness scale. J Clin Psychol 51:768–774.

Paulus MP, Feinstein JS, Leland D, Simmons AN (2005) Superior temporal gyrus and insula provide response and outcome-dependent information during assessment and action selection in a decision-making situation. NeuroImage 25:607–615.

Paulus MP, Yu AJ (2012) Emotion and decision-making: affect-driven belief systems in anxiety and depression. Trends Cogn Sci 16:476–483.

Pessoa L, Pereira M (2013) Cognition-Emotion Interactions: A Review of the Functional Magnetic Resonance Imaging Literature. In: Handbook of Cognition and Emotion (Robinson MD, Watkins ER, Harmon-Jones E, eds), pp 55-68. New York, NY, US: Guilford Press.

Plassmann H, O'Doherty JP, Rangel A (2010) Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. J Neurosci Off J Soc Neurosci 30:10799–10808.

Poldrack RA (2011) Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. Neuron 72:692–697.

Preuschoff K, Quartz SR, Bossaerts P (2008) Human Insula Activation Reflects Risk Prediction Errors As Well As Risk. J Neurosci 28:2745–2752.

Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P (1999) SENSE: sensitivity encoding for fast MRI. Magn Reson Med Off J Soc Magn Reson Med Soc Magn Reson Med 42:952–962.

Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. Nat Rev Neurosci 9:545–556.

Rangel A, Hare T (2010) Neural computations associated with goal-directed choice. Curr Opin Neurobiol 20:262–270.

Rick S, Loewenstein GF (2008) The Role of Emotion in Economic Behavior. In: Handbook of Emotions, 3rd ed. Guilford Press. Available at: http://dx.doi.org/10.2139/ssrn.954862.

Robinson OJ, Letkiewicz AM, Overstreet C, Ernst M, Grillon C (2011) The effect of induced anxiety on cognition: threat of shock enhances aversive processing in healthy individuals. Cogn Affect Behav Neurosci 11:217–227.

Robinson OJ, Overstreet C, Charney DR, Vytal K, Grillon C (2013a) Stress increases aversive prediction error signal in the ventral striatum. Proc Natl Acad Sci U S A 110:4129–4133.

Robinson OJ, Vytal K, Cornwell BR, Grillon C (2013b) The impact of anxiety upon cognition: perspectives from human threat of shock studies. Front Hum Neurosci 7:203.

Rushworth MF, Kolling N, Sallet J, Mars RB (2012) Valuation and decision-making in frontal cortex: one or many serial or parallel systems? Curr Opin Neurobiol 22(6):946-55

Schmitz A, Grillon C (2012) Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test). Nat Protoc 7:527–532.

Schwabe L, Tegenthoff M, Höffken O, Wolf OT (2012) Simultaneous Glucocorticoid and Noradrenergic Activity Disrupts the Neural Basis of Goal-Directed Action in the Human Brain. J Neurosci 32:10146–10155.

Sescousse G, Caldú X, Segura B, Dreher J-C (2013) Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. Neurosci Biobehav Rev 37:681–696.

Sharp C, Monterosso J, Montague PR (2012) Neuroeconomics: a bridge for translational research. Biol Psychiatry 72:87–92.

Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD (2004) Empathy for Pain Involves the Affective but not Sensory Components of Pain. Science 303:1157–1162.

Sokol-Hessner P, Camerer CF, Phelps EA (2013) Emotion regulation reduces loss aversion and decreases amygdala responses to losses. Soc Cogn Affect Neurosci 8:341–350.

Sokol-Hessner P, Hsu M, Curley NG, Delgado MR, Camerer CF, Phelps EA (2009) Thinking like a trader selectively reduces individuals' loss aversion. Proc Natl Acad Sci 106:5035–5040.

Sokol-Hessner P, Hutcherson C, Hare T, Rangel A (2012) Decision value computation in DLPFC and VMPFC adjusts to the available decision time. Eur J Neurosci 35:1065–1074.

Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA (1983) Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press.

Sridharan D, Levitin DJ, Menon V (2008) A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. Proc Natl Acad Sci 105:12569–12574.

Strait CE, Blanchard TC, Hayden BY (2014) Reward Value Comparison via Mutual Inhibition in Ventromedial Prefrontal Cortex. Neuron 82:1357–1366.

Talmi D, Dayan P, Kiebel SJ, Frith CD, Dolan RJ (2009) How Humans Integrate the Prospects of Pain and Reward during Choice. J Neurosci 29:14617–14626.

Tanaka SC, Yamada K, Yoneda H, Ohtake F (2014) Neural mechanisms of gain-loss asymmetry in temporal discounting. J Neurosci Off J Soc Neurosci 34:5595–5602.

Thaler RH, Johnson EJ (1990) Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. Manag Sci 36:643–660.

Tom SM, Fox CR, Trepel C, Poldrack RA (2007) The Neural Basis of Loss Aversion in Decision-Making Under Risk. Science 315:515–518.

Tracey I, Mantyh PW (2007) The cerebral signature for pain perception and its modulation. Neuron 55:377–391.

Treadway MT, Zald DH (2011) Reconsidering anhedonia in depression: lessons from translational neuroscience. Neurosci Biobehav Rev 35:537–555.

Uddin LQ, Kinnison J, Pessoa L, Anderson ML (2014) Beyond the tripartite cognition-emotion-interoception model of the human insular cortex. J Cogn Neurosci 26:16–27.

Weber EU, Blais A-R, Betz NE (2002) A Domain-Specific Risk-Attitude Scale: Measuring Risk Perceptions and Risk Behaviors. Rochester, NY: Social Science Research

Network. Available at: http://papers.ssrn.com/abstract=1422152 [Accessed July 15, 2013].

Winecoff A, Clithero JA, Carter RM, Bergman SR, Wang L, Huettel SA (2013) Ventromedial Prefrontal Cortex Encodes Emotional Value. J Neurosci 33:11032–11039.

Appendix

C The neural circuitry of emotion-induced distortions of trust

Jan B. Engelmann[1,2], Friederike Meyer[1], Christian C. Ruff[1*] & Ernst Fehr[1*]

[1]Laboratory for Social and Neural Systems Research (SNS-Lab), Department of Economics, University of Zurich, Zurich, Switzerland
[2]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

* Shared last authorship

C.1 Summary

Aversive emotions may be a key source of irrational human decision-making but still little is known about the underlying neural circuitry. Here, we show that aversive emotions distort trust decisions and cause significant changes in the associated neural circuitry. They reduce trust and suppress trust-specific activity in left temporoparietal junction (TPJ). In addition, aversive emotions reduce the functional connectivity between TPJ and emotion-related regions such as the amygdala. We also find that the posterior superior temporal sulcus (pSTS) plays a key role in mediating the impact of aversive emotions on brain-behavior relationships. Functional connectivity of right pSTS with left TPJ not only predicts mean trust taking in the absence of negative emotions, but aversive emotions also largely remove this association between TPJ-pSTS connectivity and behavioral trust. These findings may be useful for a better understanding of the neural circuitry of affective distortions and may thus help identify the neural bases of psychiatric diseases that are associated with emotion-related psychological and behavioral dysfunctions.

Highlights

- Aversive emotions are identified as an important causal factor that distorts human trust.
- The behavioral effect of aversive emotions is associated with specific distortions in the neural circuitry related to trust taking and emotion-processing.
- Aversive emotions suppress trust-specific activity in temporoparietal junction (TPJ) and reduce the functional connectivity between TPJ and amygdala.
- In the absence of aversive emotions posterior superior temporal sulcus mediates brain-behavior relationships but aversive emotions severely distort this mediation.

C.2 Introduction

Trust pervades almost every aspect of human social life. It plays a decisive role in families, organizations, markets and in the political sphere. Without trust, families fall apart, organizations are inefficient, market transactions are costly and political leaders lack public

support. Recent research in behavioral economics and neuroeconomics has begun to elucidate the determinants and neural correlates of trust (Bohnet and Zeckhauser, 2004; Delgado et al., 2005). However, despite recent progress in understanding the determinants of trust (Fehr, 2009) and its distortions in psychiatric disorders (Gromann et al., 2013; King-Casas et al., 2008), relatively little is known about the impact of our emotions on trust taking. Emotions, in particular those with high intensity, can have deleterious effects on our decision-making faculties, as hinted at by a multitude of public press reports and recent theoretical (Loewenstein, 2000) and experimental accounts (Knutson et al., 2008; Kuhnen and Knutson, 2011; Phelps, 2009; Schulreich et al., 2014). It is therefore important to understand the behavioral and neural mechanisms by which emotions distort decisions to trust.

While much progress has been made in outlining the neural underpinnings of emotional processes on the one hand (Lindquist et al., 2012; Pessoa and Adolphs, 2010) and of decision-making on the other (Kable and Glimcher, 2009; Rangel and Hare, 2010), the effects of emotion on choice have received relatively little attention to date (Kuhnen and Knutson, 2011; Phelps, 2009; Schulreich et al., 2014). Theoretical accounts of the influence of emotion on choice (Loewenstein, 2000) distinguish between two types of emotions: *anticipatory emotions*, such as the anticipated pleasure from the future consumption of a good that reflect how decision-makers expect to feel about the outcomes of their decisions, and *incidental emotions* that occur at the time of the decision, but are unrelated to the choice outcomes. Incidental emotions are of particular interest because of their ubiquity in real life and because they are prime candidates for emotion-induced behavioral distortions. By definition, incidental emotions are unrelated to choice outcomes and, to the extent to which they affect behavior, may cause irrational behavioral biases.

To study the behavioral impact and the underlying neural circuitry of affect-induced distortions of trust, we adapted the trust game (Berg et al., 1995) to an imaging context. In the trust game, two anonymous players, which we call investor and trustee, sequentially send money to each other. In the first stage, the investor faces the choice of whether and how much of her endowment to transfer to the trustee. Then the experimenter triples the sent amount, before it is transferred to the trustee. The investor's decision to transfer money thus increases the total amount of money that can be distributed among the two players. In the second stage, the trustee is informed about the total amount that he received after which he can send back part or all of this money. Thus, while the investor's transfer increases the total amount of money available to both parties, the investor also faces the risk of

benefitting nothing from the transfer because the trustee is completely free in his back-transfer decision. Therefore, the decision to transfer money constitutes an act of trust, as the investor makes herself vulnerable to the potentially selfish behavior of the trustee (Fehr, 2009).

Such trust taking involves both a financial risk due to the possibility of losing the invested money, as well as a social risk of being betrayed by an untrustworthy trustee. Therefore, to enable clear identification of the impact of incidental emotion on the mechanisms involved in the social aspects of trust taking, it is important to include a well-matched non-social control task. For this reason, our subjects also faced a non-social control condition that was identical to the trust condition in every respect except that instead of a trustee, a computer made a "back-transfer" that determined the profitability of the investor's "transfer" (Baumgartner et al., 2008; Kosfeld et al., 2005). The profitability of the investor's transfer in the non-social control and trust condition was exactly the same because the investors had exactly the same choice options in both conditions and the computer sampled the "back-transfer" decisions in the non-social control condition according to the probability distribution of back-transfers that was generated by the trustees' decisions in the trust condition (see experimental procedures).

Subjects made decisions in either trust or non-social control trials within two different emotional contexts. They were either under the threat of relatively intense tactile stimulation that was somewhat painful ("threat condition"), or they faced the possibility of receiving weak tactile stimulation in the "no threat" condition (Figure C-1). A prolonged period of aversive affect was established by administering the tactile stimulations at unpredictable time points and frequencies for the duration of an entire block. A block consisted of several trust or control trials in the threat condition or several trials in the no-threat condition. This approach has a number of advantages over typical emotion-induction techniques that have been employed in previous investigations of the influences of emotion on decision-making (Harlé et al., 2012; Lerner et al., 2013). The limitations of standard emotion induction procedures have been well-documented in the past (Martin, 1990; Westermann et al., 1996) and include susceptibility to demand effects (Buchwald et al., 1981; Polivy and Doyle, 1980), emotional responses of insufficient intensity (Marston et al., 1984; Martin, 1990), systematic differences between subjects in susceptibility to the induction procedure (Blackburn et al., 1990), as well as a relatively brief time course of the effect, the duration of which may vary by subject and may not suffice for experiments longer than 10 minutes (Frost and Green, 1982; Isen et al., 1976). The threat-of-shock

paradigm employed in the current study has been shown to reliably induce negative affect (Bogdan and Pizzagalli, 2006; Grillon et al., 1993; Schmitz and Grillon, 2012) and addresses the limitations of standard emotion induction procedures as follows: (1) threat of shock provides an immediate stimulus of biological significance that triggers an aversive and automatic emotional reaction, the intensity of which can be measured throughout the experiment using standard psychophysiological techniques; (2) using a blocked threat-of-shock paradigm reinstates the emotional reaction at every presentation, which can thus be maintained for the duration of the entire experiment; (3) threat of shock was administered within-subject, therefore allowing each subject to serve as their own control; (4) tactile stimulation was administered in both the trust and the non-social control condition, thus minimizing demand effects.

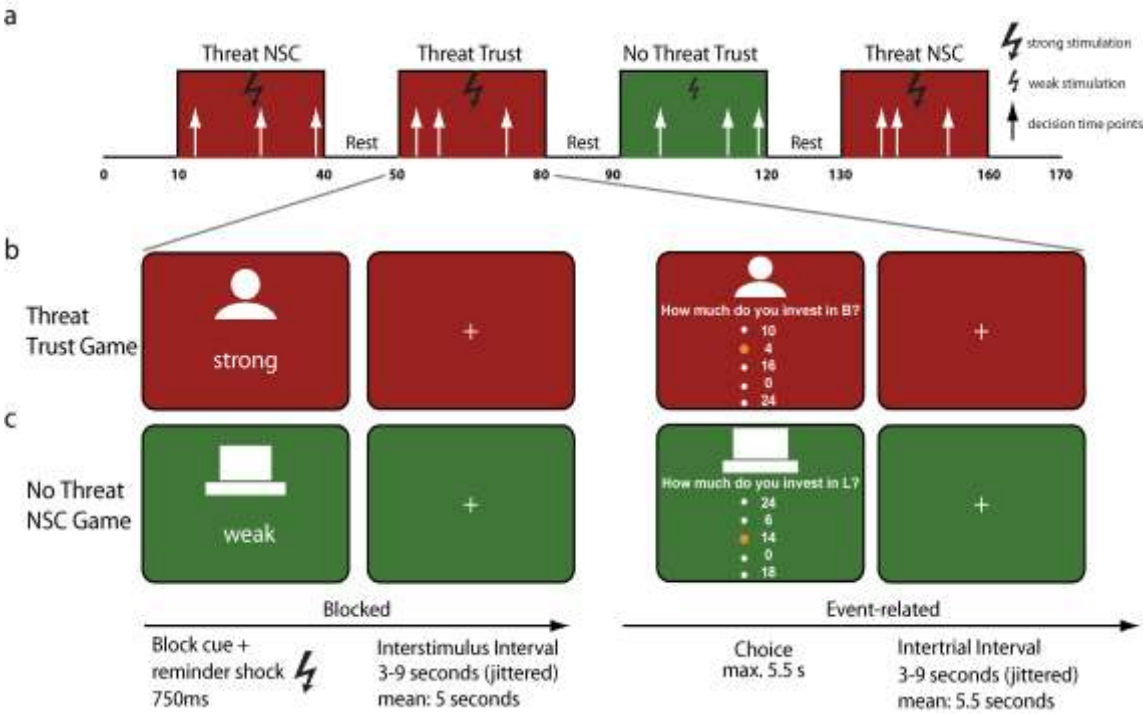

Figure C-1: Schematic representation of hybrid fMRI design, trial sequence and timing. (a) Subjects faced blocks of trust and non-social control (NSC) trials in random order. During a trust block they played 3 trust games with different anonymous partners. During an NSC block they played 3 non-social control (NSC) games in which a pre-programmed computer chose "back transfers". Both during trust and NSC blocks subjects expected either strong ("threat") or weak ("no-threat") tactile stimulation at unpredictable times during the block. (b) Sequence and timing of a trial in a trust block with strong tactile stimulation ("threat"). (c) Sequence and timing of a trial in a non-social control block with weak tactile stimulation ("no-threat"). At the beginning of each block, a 750-ms visual cue reminded subjects of the game type (trust or non-social control) and stimulation intensity (weak or strong) for the current block. The "human" icon indicated a trust block, the "computer" icon represented a control block. Two additional cues indicated the threat level to the subjects: (i) an actual tactile stimulation ("reminder shock") at the end of the block cue indicated the current block's stimulation intensity level and (ii) the background color informed the subjects throughout the entire block whether they have to expect a strong

or weak stimulation (in Figure C-1, e.g., red = strong and green = weak shock, stimulation intensity-color association was counterbalanced). After a jittered interstimulus interval of three to nine seconds, the first of three trials was presented. In a trust game trial, subjects chose how much of their endowment of 24 CHF to transfer to a stranger ("How much do you invest in B?"). In a non-social control game trial, subjects chose how much of their endowment to invest in an ambiguous lottery that provided a 40-60% probability of returning an amount greater than the investment ("How much do you invest in L?"). To ensure subjects' attention, the discrete investment opportunities – in the range between 0 and 24 CHF – randomly changed from trial to trial.

Previous fMRI investigations of emotion (Lindquist et al., 2012) and cognitive control strategies through which emotional reactions can be modified (Ochsner et al., 2012) inform the neurobiological hypotheses of the current investigation. We conjectured that incidental affect modulates trust-specific computations, such as the assessment of the trustee's trustworthiness. Neurally, we expected, therefore, that aversive affect may influence trust decisions by specifically modulating responses in regions involved in representing other people's mental states, including temporoparietal junction and dorsomedial prefrontal cortex (Mitchell et al., 2006; Saxe and Kanwisher, 2003; Young et al., 2010). Given our manipulation of affective state, we also expected regions known to play a role in emotion processing to be involved in the influence of aversive affect on trust decisions. Such regions include the amygdala and anterior insula, which have consistently been implicated in processing negative emotion (LeDoux, 2007; Wicker et al., 2003), as well as ventral striatum and ventromedial prefrontal cortex, which have consistently been implicated in processing the rewarding aspects of stimuli (Grabenhorst and Rolls, 2011; Schultz et al., 1997). Notably, recent meta-analyses have underlined the importance of interacting networks in the production (Lindquist et al., 2012) and regulation of emotions (Ochsner et al., 2012). The evidence suggests that the experience and regulation of emotions emerges through the interaction of various nodes that include areas involved in visual perception, emotion, as well as executive mechanisms, including dorsolateral PFC (Hariri et al., 2003), anterior cingulate cortex (Ochsner et al., 2004; Shackman et al., 2011), and ventrolateral PFC (Wager et al., 2008). We therefore conducted functional connectivity analyses to investigate the impact of aversive affect on interacting networks involved in trust taking.

C.3 Results

Electrophysiological results

We scanned 41 volunteers while they made trust decisions during the emotionally aversive threat condition and during the emotionally neutral no-threat condition. Emotional arousal during the threat and no-threat condition was assessed using galvanic skin conductance responses (SCR). Mean SCRs were extracted for 17 one-second time bins following trial onset using an approach commonly employed in analyses of fMRI data, as outlined in detail in experimental procedures. As illustrated in Figures C-2a and C-2b, mean SCR during both trust and non-social control trials were significantly greater during the threat condition compared to the no-threat condition. This observation was confirmed by three-way repeated measures ANOVA of SCRs with the factors time (0-16 seconds post trial onset), choice domain (trust, control), and threat (absent, present). We found significant main effects of threat [$F(1,39) = 141.3$, $p < 0.001$, $\eta^2 = 0.784$] and time [$F(16,624) = 12.12$, $p < 0.001$, $\eta^2 = 0.237$] and a significant interaction between threat and time [$F(16,624) = 99.28$, $p < 0.001$, $\eta^2 = 0.718$], indicating that the impact of threat relative to no threat on SCR differed across time. However, we did not obtain significant interactions between threat and choice domain [$F(1,39) = 0.006$, $p = 0.938$, $\eta^2 < 0.001$], and between time and choice domain [$F(16,624) = 0.905$, $p = 0.56$, $\eta^2 = 0.023$] indicating that threat had similar impacts on SCRs for both the trust and non-social control task. To further specify the effects of threat on SCR during trust and control trials, we performed follow-up pairwise comparisons at each time point, correcting for multiple comparisons across time-bins. Significantly enhanced SCR during threat relative to no threat were observed from 2 until 16 seconds post trial onset during trust decisions and from 3 until 16 seconds post trial onset during non-social control trials (all $t(39) > 3.313$, all $p < 0.002$). Taken together, these results indicate significantly greater emotional arousal during the threat condition relative to the no-threat condition in both choice domains.

Manipulation Checks

The emotional arousal illustrated in Figures C-2a and C-2b was clearly experienced as aversive by the subjects. In an open-ended questionnaire administered after scanning, 95.12% of subjects responded that they experienced aversive emotional arousal during threat blocks (Figure S1a). The aversive nature of the threat condition was further

confirmed by strong activations of central nodes of the brain's pain matrix during strong compared to weak tactile stimulation (Figure S1b) and by the observation of enhanced SCRs following strong tactile stimulation compared to weak tactile stimulation (Figure C-2c).

The above electrophysiological results, in combination with the self-reported emotions and the activation of the brain's pain matrix during and after the shock, indicate that subjects experienced the threat of a shock as an aversive and arousing emotional state. This state is clearly unrelated to the monetary outcome of trust- and risk-taking, as it does not affect the trustee's or the computer's decisions. The next question we addressed is whether this incidental emotional state distorts subjects' behavior relative to the no-threat control condition.

Behavioral results

To identify whether aversive affect had a significant impact on decision-making, we computed mean transfer rates and mean reaction times during trust and non-social control trials for each emotional context and submitted these data to a two-way repeated-measures ANOVA with the factors choice domain (trust, control) and threat (absent, present). Aversive affect significantly reduced transfers during both trust and non-social control trials in a similar fashion (Figure C-2d), as indicated by significant main effects of choice domain [$F(1,40) = 20.319$, $p < 0.001$, $\eta^2 = 0.337$] and threat [$F(1,40) = 17.483$, $p < 0.001$, $\eta^2 = 0.304$] but no interaction [$F(1,40) = 0.122$, $p = 0.7$, $\eta^2 = 0.003$]. Moreover, separate pairwise comparisons showed that the threat condition significantly reduced mean transfers (Figure C-2d) in the trust game [$t(40) = -3.4$, $p < 0.005$, mean transfer difference = -1.1 CHF] and in the non-social control game [$t(40) = -3.16$, $p < 0.005$, mean transfer difference = -0.93 CHF]. To exclude the possibility that choices were affected by the *actual experience* of shocks, rather than by the ongoing aversive affective context due to shock expectation, we ran several multiple regression analyses (Text S1). The regression results (Table S1) show that the behavioral results reported above were indeed due to the aversive affective context ($p < 0.001$) generated by the *threat* of a shock, rather than reflecting the effect of actual shock experience immediately before decisions are taken ($p = 0.23$).

Aversive affect also led to faster reaction times during both trust- and non-social control trials (Figure C-2e). We obtained significant main effects of choice domain [$F(1,40) = 6.226$, $p = 0.017$, $\eta^2 = 0.135$] and threat [$F(1,40) = 17.01$, $p < 0.001$, $\eta^2 = 0.298$] but no interaction [$F(1,40) < 0.001$, $p = 0.993$, $\eta^2 < 0.001$]. The main effect of threat is

characterized by significantly faster mean reaction times in the threat relative to the no-threat condition (Figure C-2e) for both the trust game [$t(40) = -3.3$, $p < 0.005$, mean RT difference = -0.13 s] and the non-social control task [$t(40) = -2.5$, $p < 0.05$, mean RT difference = -0.13 s].

Taken together, these behavioral results indicate that aversive affect significantly reduced transfer rates and reaction times in both the trust and the nonsocial control task. Notably, the absence of a significant interaction between threat and choice domain for a whole range of non-neural measures (transfer rates, response latencies, SCR) indicates that the impact of aversive affect during trust and nonsocial control trials is similar across multiple measurement modalities, indicating that our non-social condition constitutes a well-matched control for the trust game.

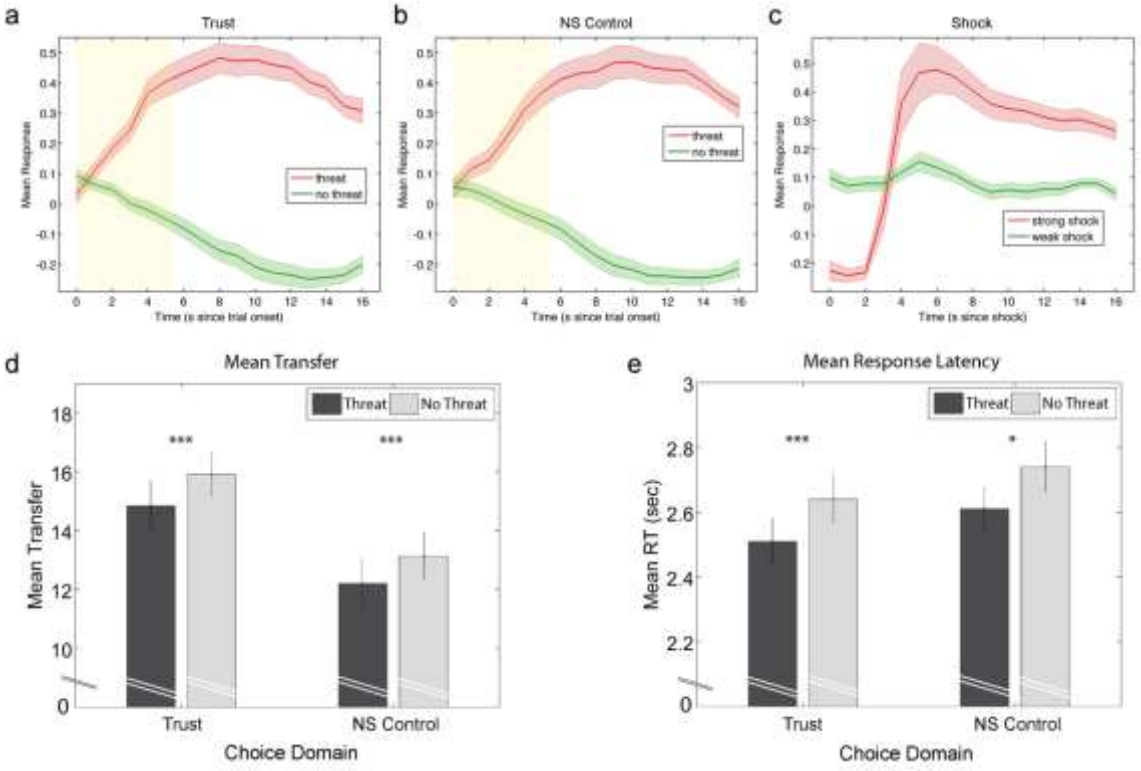

Figure C-2: Psychophysiological and behavioral results. Aversive affect (threat condition) significantly influences emotional arousal, mean transfer rates, and reaction times. The threat of an aversive tactile stimulation leads to a strong increase in skin conductance responses (SCR) in (a) the trust game [main effect of threat: $F(1,39) = 90.74$, $p < 0.0001$, $\eta^2 = 0.699$] and (b) the non-social control game [main effect of threat: $F(1,39) = 110.95$, $p < 0.0001$, $\eta^2 = 0.74$]. Note that these increases in SCR's reflect generalized arousal during the decision-period [5.5 sec., shown as yellow background in (a) and (b)] due to the *threat* of strong tactile stimulation and are independent of the shock itself. The SCR effect of the shock itself is illustrated in (c), which shows the SCR after an *actually experienced* strong tactile stimulation (red line). The actual experience of a weak tactile stimulation leaves SCR almost unchanged (green line). (d) In the threat condition (relative to the no-threat condition) subjects transferred significantly less to an anonymous stranger in the trust game [$t(40) = -3.4$, $p < 0.005$] and invested less

into an ambiguous lottery in the non-social control game [$t(40) = -3.16$, $p < 0.005$]. These results are driven by the emotional arousal induced by the threat of a shock and not by the actual experience of shocks shortly before choice (Table S1). (e) In the threat condition (relative to the no-threat condition) subjects made their decisions significantly faster in both the trust [$t(40) = -3.3$, $p < 0.005$] and the control [$t(40) = -2.5$, $p < 0.05$] game. *** $p < 0.005$; ** $p < 0.01$; * $p < 0.05$.

fMRI results

The main goal of our fMRI analyses was to identify the neural circuitry underlying affect-induced distortions of trust decisions. We therefore first examined brain activation in the ROIs that we conjectured (see our hypotheses in the introductory section) to be preferentially engaged during trust decisions. For all these analyses, we employed small-volume correction at an FWE-corrected threshold of $p < 0.05$ *in truly independent ROIs* defined with reverse inference maps from relevant search terms on neurosynth.org (Yarkoni et al., 2011) (see experimental procedures). As a first step, our analyses confirmed that several of the conjectured regions were indeed specifically involved in decisions during trust (vs. non-social control) trials when aversive affect was absent. That is, areas in left TPJ (-57, -60, 27; k = 247; red region in Figure C-3a), dorsomedial prefrontal cortex (-9, 60, 18; k = 39; red region in PFC in Figure C-3a), and ventromedial prefrontal cortex (9, 32, -12; k = 89) showed significantly greater activation during decision-making in trust relative to control trials in the absence of threat (all activations SV FWE-corrected, see also Table S2a). We then examined which of the ROIs showed a breakdown of trust-specific activity due to aversive affect by computing the following contrast: $[(\text{Trust}_{\text{No threat}} > \text{Trust}_{\text{Threat}}) > (\text{NS Control}_{\text{No threat}} > \text{NS Control}_{\text{Threat}})]$. A region in left TPJ indeed showed a significant interaction effect (-60, -54, 19; k = 95, SV FWE-corrected, yellow region in Figure C-3a, see also Table S2b). To further characterize this interaction effect we examined post hoc the impact of aversive affect in the trust game separately from its impact in the non-social control game: the results of the contrast (No Threat > Threat) during trust trials shows a suppression of activation for trust decisions within left TPJ (-58, -55, 19; k = 103; SV FWE-corrected, Table S2c). During non-social control decisions on the other hand, no voxels in our left TPJ ROI showed greater activity during no threat relative to threat, even at a very liberal threshold of $p < 0.05$, uncorrected (see also Figure S2a for additional analyses that underline the strength of the interaction effect in left TPJ). These results indicate that the interaction effect is based on a selective interference of aversive affect with trust-related activity, but not with activity in the nonsocial control condition.

Given our hypothesis that decisions involving trust rely on neural circuitry that

mediates the interplay between social cognition and valuation, we also examined the impact of aversive affect in other a-priori regions of interest commonly implicated in reward processing (vmPFC and ventral striatum) and emotion processing (bilateral amygdala). These results show reductions in trust-related activity due to aversive affect in vmPFC and ventral striatum (Table S2c). For completeness, we also conducted a whole brain analysis to identify potentially important activations outside our a-priori regions of interest. The whole brain analysis shows that in the left superior temporal sulcus, posterior cingulate cortex and left inferior parietal lobe (Figure S3 and Table S3) aversive affect suppresses brain activity only in the trust game but not in the nonsocial control game. However, in these regions we do not find a significant interaction effect in the sense that suppression of activity due to aversive affect is significantly stronger in the trust relative to the nonsocial control task.

The network effects of aversive affect during trust decisions

Recent studies stress the importance of the interplay between cognitive and emotional networks (Lindquist et al., 2012). Therefore, we investigated the effects of aversive affect on the connectivity between trust-relevant brain regions with Psychophysiological Interaction analyses (McLaren et al., 2012). In view of the key role of the temporoparietal junction (TPJ) in perspective taking and mentalizing (Van Overwalle, 2009; Young et al., 2010), the conjecture that these mental operations are important for trust taking, and our finding of enhanced activity in this area in the trust compared to the control task (see above), we were particularly interested in how aversive affect changes the functional connectivity between the TPJ and emotion processing regions, such as the amygdala. ROI analysis of the data during the *threat-absent* condition confirmed that the connectivity between the TPJ and a region in left amygdala (-22, -9, -15; $p < 0.05$, SV FWE-corrected, $k = 14$) is indeed stronger during decisions in the trust relative to the control task (Table S4a). Next, we were interested whether the threat-induced connectivity change is larger in the trust condition than in the control condition. For this purpose we conducted an independent interaction analysis of the connectivities in the trust and the control condition via the contrast: $[(Trust_{No\ threat} > Trust_{Threat}) > (Control_{No\ threat} > Control_{Threat})]$. This analysis revealed threat-induced aversive affect causes a stronger connectivity change between TPJ and a region in the amygdala (-26, 0, -23; $p < 0.05$, SV FWE-corrected, $k = 29$, Figure C-3c shown in red) in the trust game compared to the control task (see Table S4b). We performed a post-hoc inspection of the significant interaction in left amygdala by investigating the effect of threat on connectivity changes for the trust and the non-social control condition

separately. We find that aversive affect disrupted functional connectivity specifically during trust (compare red vs green bars in Figure C-3d) but not during non-social control decisions. A follow-up contrast of No threat > Threat during trust decision confirmed the suppression of TPJ-amygdala connectivity during trust decisions (-28, -6, -14; $p < 0.05$, SV FWE-corrected, $k = 110$, Table S4c). In contrast, during decisions in the non-social control task no voxels in our left amygdala ROI showed greater connectivity during no threat relative to threat, even at a very liberal threshold of $p < 0.05$, uncorrected (see also Figure S2b for additional analyses that underline the strength of the interaction effect in left amygdala). Thus, aversive affect not only affected trust-specific overall activation in the TPJ, but also led to trust-specific connectivity changes of this area with the amygdala.

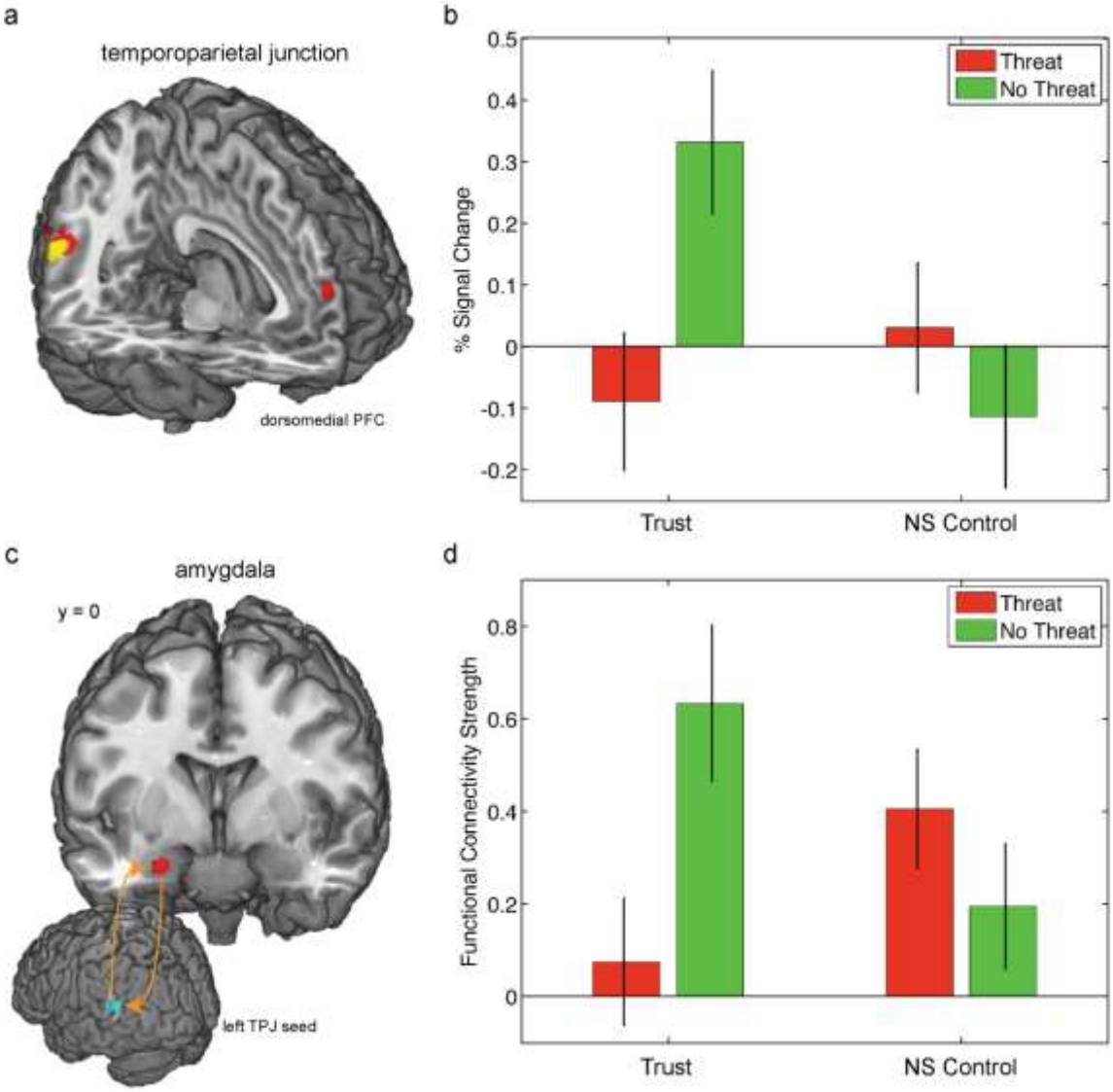

Figure C-3: The impact of aversive affect on trust-specific activity and connectivity in the temporoparietal junction (TPJ). (a) Preferential involvement of the left temporoparietal junction (-57, -60, 27; $p < 0.05$, SV FWE corrected, region shown in red, $k = 247$;) as well as the dorsomedial (-9, 60, 18; $k = 39$; $p < 0.05$, SV FWE corrected, region in PFC shown in red) and ventromedial PFC (9, 32, -12; $k = 89$; $p < 0.05$, SV FWE corrected, not shown) in trust decisions when aversive affect is absent (see Table S2a). Importantly, the threat of a shock reduced activation in left TPJ (relative to "no threat") significantly more during the trust game compared to the nonsocial control game (-60, -54, 19; $p < 0.05$, SV FWE corrected, $k = 95$, see Table S2b). Voxels whose activity reflects this interaction effect are shown in yellow. (b) Aversive affect significantly suppresses activity in left TPJ only during decisions in the trust game (-58, -55, 19; $p < 0.05$, SV FWE corrected, $k = 103$, Table S2c), but not in the non-social (NS) control game. (c) The left amygdala shows both significantly greater connectivity with TPJ during trust relative to control when aversive affect is absent (-22, -9, -15; $p < 0.05$, SV FWE corrected, $k = 14$, Table S4a) and a significant interaction effect (red, -26, 0, -23; $p < 0.05$, SV FWE corrected, $k = 29$, Table S4b). (d) The interaction is characterized by a significant suppression of functional connectivity between the left TPJ and left amygdala during decisions in the trust game (-28, -6, -14; $p < 0.05$, SV FWE-corrected, $k = 110$, Table S4c), while no suppression is observed in the non-social control task. The follow-up statistical tests were conducted via simple effects contrasts in neuroimaging space testing activation (or connectivity) suppression due to aversive affect during trust and non-social control decisions separately. Activation patterns illustrated in the bar charts reflect signal (b) and connectivity (d) change within a 6 mm sphere around the interaction peak voxel in left TPJ (b) and left amygdala (d).

The above PPI analyses show the average impact of aversive affect on the functional connectivity between TPJ and amygdala. However, as we observed strong individual differences in the functional connectivity between TPJ and amygdala on the one hand, and in mean transfer levels on the other hand, we next asked the question how individual differences in functional TPJ connectivity are related to individuals' mean transfer levels in the absence and the presence of aversive affect. For this purpose, we first examined whether in our a priori regions of interest the relationship between mean transfers and the functional TPJ connectivity is different in the trust game compared to the control game (in the absence of threat). In the conjectured ROIs, only the right amygdala [27, 2, -21, k = 48, Table S5] showed such a pattern. In the next step we therefore conducted an exploratory whole-brain analysis that identified an extended network of regions (Figure C-4 and Table S6a) that showed a difference in the relationship between mean transfers and the functional TPJ connectivity across trust and control tasks (in the absence of threat). The regions in this network comprised dorsomedial PFC [-3, 18, 55, k = 5069], superior temporal sulcus extending into angular gyrus [63,-45,6, k = 1076], bilateral ventrolateral PFC [left: -50, 23, -8, k = 695; right: 57, 20, 10, k = 544], intraparietal sulcus [46, -58, 45, k = 532] and dorsolateral PFC [50, 12, 37, k = 1464]. In all these regions we observe a positive and significantly stronger correlation between mean transfer levels and functional TPJ connectivity in the trust compared to the control task (see Figure C-4 and Table S6a). Taken together, the above results therefore confirm the conjecture that there is specific functional connectivity between the TPJ and amygdala that relates to behavioral trust and that this specific relationship between TPJ connectivity and trust extends to a network consisting of STS, dmPFC and bilateral vlPFC.

How does the relationship between functional connectivity patterns in the trust network and behavioral trust-taking change if subjects are exposed to aversive affect? To answer this question, we examined whether aversive affect caused significant changes in the relationship between TPJ connectivity (with its target regions) and mean trust levels in the trust game. This analysis showed that aversive affect caused a general breakdown of the association between left TPJ connectivity and mean trust taking in the right posterior superior temporal sulcus [64, -43, 4, k = 514, Table S6b]. In particular, there is a significantly positive relationship between left TPJ connectivity and mean trust levels during the no-threat condition (Figure C-4d, green regression line) that vanishes in the presence of threat (Figure C-4d, red regression line). Please note that the region in pSTS that is affected in this way by aversive emotions significantly overlaps with the region that

shows stronger TPJ connectivity in the trust relative to the control task (region shown in yellow in Figure C-4). These results suggest that connectivity between left TPJ and its target region in pSTS supports general trust taking when distortionary aversive affect is absent. However, in the presence of aversive affect the relationship between the connectivity pattern in the trust network and behavioral trust taking reverses. Thus, aversive affect not only reduces average trust taking, but it also diminishes specific relationships between the connectivity patterns in the trust network and behavioral trust taking. Our results therefore suggest that the pSTS is a crucial neural node that mediates the breakdown of trust in the presence of threat.

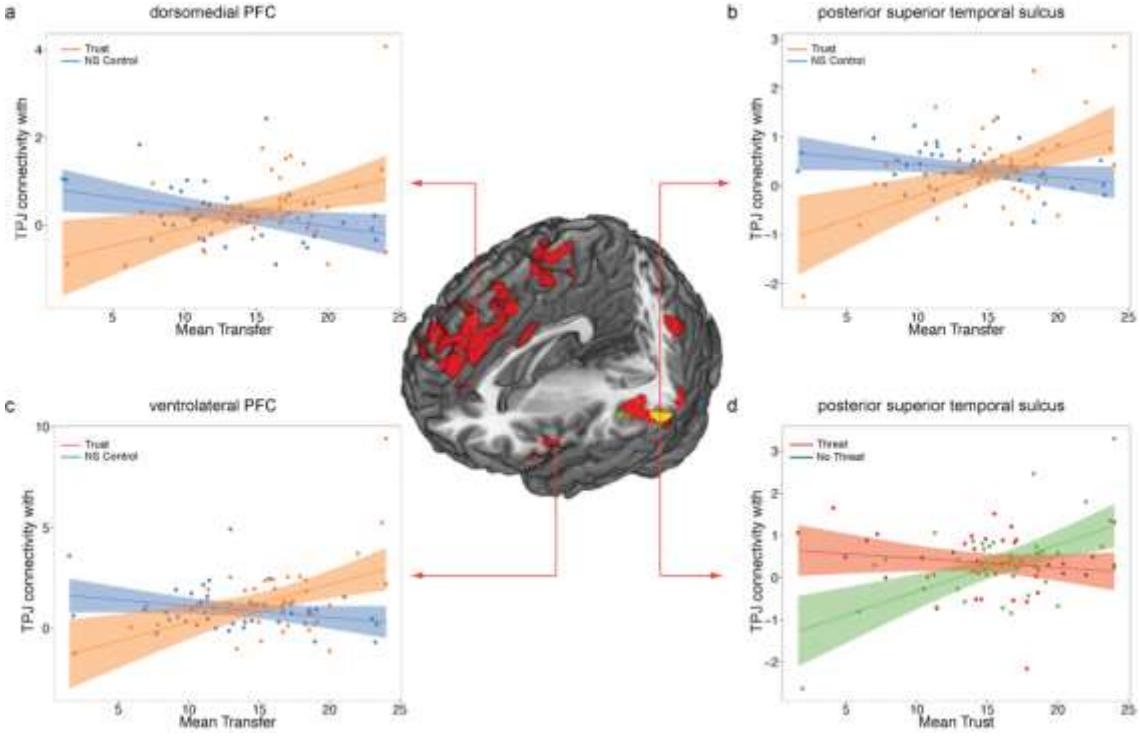

Figure C-4: Trust-specific functional connectivity (a-c) and threat-induced breakdown of connectivity (d) between TPJ and a network of target regions. (a-c) Connectivity between left TPJ and its targets (shown in red) is positively associated with trust taking (orange regression lines) during the no-threat condition in (a) dorsomedial PFC [-3, 18, 55, k = 5069], (b) posterior superior temporal sulcus [63, -45,6, k = 1076], (c) bilateral ventrolateral PFC [left: -50, 23, -8, k = 695; right: 57, 20, 10, k = 544]. In contrast, mean transfers (i.e. investments) during the non-social (NS) control task (blue regression lines) are associated with reduced connectivity strength between TPJ and these regions. In all cases (a-c), the correlation between mean transfer and connectivity strength is stronger in the trust game compared to the non-social control task (whole brain analysis, $p < 0.05$, FWE corrected at cluster level, see Table S6a). The intraparietal sulcus [46, -58, 45, k = 532] and dorsolateral PFC [50, 12, 37, k = 1464] show a similar pattern (see Table S6a) but are not shown in the figure. (d) Aversive affect causes a breakdown of the association between TPJ-pSTS connectivity and mean trust. The correlation between mean trust levels and TPJ-pSTS connectivity is stronger in the no threat compared to the threat condition [yellow region, 64, -43, 4; k = 514; whole brain analysis, $p < 0.05$, FWE corrected at cluster level, see Table S6b]. Specifically, there is a positive association between TPJ-pSTS connectivity and the mean trust level when distortionary aversive affect is absent (green regression line), which is eliminated by threat (red regression line). This suggests that connectivity between TPJ and its target region in pSTS supports general trust

taking only in the absence of threat. The regression lines in (a-d) predict functional connectivity strength as a function of mean transfer levels based on an extended OLS model that estimates both the slope of the relationship between mean transfers and functional connectivity in the non-social control task and the increase in this relationship in the trust task (relative to the non-social control task). For this purpose we extracted the data from 6 mm spheres around interaction peak voxels (see experimental procedures). Confidence bounds around regression lines reflect 95% confidence intervals around the model fit.

The impact of aversive affect on domain-general neural circuitry

The previous analyses indicate that aversive affect had distinct effects on neural processes devoted to decision-making in the trust task relative to the non-social control task. Given the similar *behavioral* impact of aversive affect on trust and control decisions, we further addressed whether there are any domain-general effects of aversive affect on the neural correlates of decision-making that are independent of choice-domain. We identified a domain-general network of regions that show suppression and enhancement in choice-related neural activity during both the trust and the control task (Table S7). We observed the suppression of neural activity in the threat condition (red time course, Figure S4) relative to the no threat condition (green time course, Figure S4) in bilateral posterior dlPFC [left: -62, -4, 18, k = 1901 and right: 62, -6, 28, k = 1010], left amygdala [-24, -15, -23, k = 552], posterior paracentral lobule [4 -36, 69, k = 887], and a large cluster (k = 4082) that includes left vlPFC [-48, 41, -8] and vmPFC [-10, 44, -8]. Significant enhancement of activity during decision-making under aversive affect was obtained in the thalamus [18, -6, 1, k = 559] and cerebellum [-4, -46, -24, k = 849]. Together, these results identify a network of domain-general regions, whose decision-related activity is significantly impacted by incidental aversive affect.

C.4 Discussion

Incidental affect is a ubiquitous phenomenon that pervades many aspects of human behavior and human social interaction. However, despite its importance our knowledge about the behavioral and neural effects of incidental affect is still rather limited. In this paper, we investigated the behavioral and neural impact of incidental affect on trust decisions. We employed a novel experimental technique to establish aversive affect by inducing a prolonged expectation for unpredictable and aversive tactile stimulation embedded within a hybrid fMRI design. The threat of painful tactile stimulation significantly increased autonomic arousal during both social and non-social decision-

making and was associated with consistent self-reports of the experience of aversive affect. Behaviorally, we observed that aversive affect significantly reduced transfer rates to partners in a trust game, indicating a significant reduction in participants' willingness to trust a stranger. Importantly, despite the fact that aversive affect was incidental to the decisions made by subjects, we observed significant behavioral, autonomic and neural effects of the aversive affective state. The behavioral impact of aversive affect contradicts economic models of choice (Loewenstein, 2000) and underlines the importance of emotions in decision-making, even when they are unrelated to the choice outcomes.

Our neuroimaging results provide information about the neural mechanisms behind the suppression of participants' propensity to trust. Several regions such as the temporoparietal junction, dorsomedial (dmPFC) and ventromedial prefrontal cortex (vmPFC) were preferentially engaged during trust decisions, and aversive affect led to a trust-specific breakdown of this activation pattern (Figures C-3a-b) in the left temporoparietal junction (TPJ). Moreover, in the absence of aversive affect, we observed significant connectivity between TPJ and amygdala during trust taking, but this connectivity was completely disrupted during aversive affect (Figures C-3c-d). Aversive affect also disrupted the relation between the connectivity patterns in the neural network underlying trust and behavioral levels of trust taking. In particular, functional connectivity strength with the TPJ predicted mean trust levels for a network of regions consisting of amygdala, dorsomedial PFC, superior temporal sulcus, as well as ventrolateral and dorsolateral PFC (Tables S5 and S6, Figures C-4a-c). Aversive affect caused a specific breakdown of this association for the posterior STS, such that the connectivity between left TPJ and right pSTS no longer predicted overall trust taking (Figure C-4d). Our results therefore identify a network of interconnected regions consisting of left TPJ, amygdala and right pSTS, whose connectivity patterns during trust taking are significantly impacted by aversive affect. These results provide potentially important insights into the neural circuitry that enables people to trust, and they may, therefore, generalize to many social situations that require trust. Given that psychiatric diseases, such as pathological anxiety, social phobia or depression, are characterized by a particularly pronounced susceptibility to negative affect, our results may also be useful in understanding the neural circuitry associated with affect-related distortions of social behavior in psychiatric diseases.

The temporoparietal junction (TPJ) has repeatedly been implicated in representing and interpreting others' mental states and behaviors (Behrens et al., 2009; Morishima et al., 2012; Young et al., 2010). During trust taking, humans have to form expectations about

their partners' responses to their trust, which critically involves their mentalizing faculties. Our results suggest that enhanced connectivity between TPJ and regions important for social cognition (dorsomedial PFC, superior temporal sulcus), emotion processing (amygdala) and cognitive control (vlPFC, dlPFC) supports individuals' trust taking. Aversive affect suppresses these neural mechanisms by (1) reducing the activation level in TPJ, (2) by disrupting the association between TPJ connectivity and mean trust taking, as well as (3) by reducing the connectivity between TPJ and amygdala. Together, these results suggest that aversive affect weakens neural mechanisms important for representational capacities in the social domain, as well as the transfer of these representations to emotion and cognitive control regions during decision-making, thereby leading to a decrease in trust taking.

In conclusion, we report results that show a strong behavioral impact of incidental emotions on social (trust) and non-social (control) decision-making and we identify the neural mechanisms associated with the impact of aversive affect on trust taking. These effects were observed even though induced emotions were unrelated to the choice outcomes in our task, confirming that incidental emotions can have a powerful impact on behavior and its underlying mental operations. Our findings inform the development of economic and social theory and call for the integration of incidental affect in behavioral models of social and non-social decision-making. In addition, by identifying the specific distortions of the neural circuitry supporting trust taking, we provide a first step towards neural models that help us better understand such distortions. In particular, our results support the notion that an important mechanism through which aversive incidental affect impacts social decision-making is the suppression of activity and connectivity between regions known to be crucial for social cognition, such as the temporoparietal junction, superior temporal sulcus and amygdala. We therefore tentatively suggest that one potential effect of aversive emotions may be a reduced ability to consider the consequences of one's actions for others or on other's behavior. Such enhanced self-orientation during heightened aversive emotional arousal, such as anger or the threat of negative outcomes, has perhaps an evolutionary origin that may have served survival during social conflict situations. In the context of the modern world, however, such emotional reactions can cloud our judgments and thereby lead to maladaptive behaviors.

C.5 Experimental procedures

Participants

41 right-handed volunteers (mean age (std.) = 22 (2.145), 17 females) from various Universities in Zurich participated in the current experiment. Participants gave written informed consent to procedures approved by the local ethics committee (Zurich, Switzerland) before participating in the study. Subjects were right-handed as assessed by the Edinburgh handedness questionnaire and did not report any history of psychological illness or neurological disorders, as assessed by a standard screening form.

Prescanning procedure

Particular care was taken to ensure that subjects understood all aspects of the experiment. To this end, subjects were instructed to carefully read detailed instructions and were required to fill out an extensive questionnaire probing their understanding of the experimental procedures. The accuracy of each subject's answers was confirmed by the experimenters and discussed in a brief interview that lasted for ca. 10 minutes. Subjects were then placed inside the scanner for a brief practice session consisting of 12 trials to ensure that they could view all stimuli, perform the task, make decisions in the allotted 5.5 seconds per trial, understood the experimental setup and to give subjects the opportunity to ask further questions.

After completion of practice, subjects were taken out of the scanner and washed their hands before placement of SCR and stimulation electrodes. Subjects were then placed inside the scanner and two ring electrodes were attached to the dorsum of the left hand: (1) the electrode providing relatively higher intensity stimulation was placed between one to two cm below the second carpometacarpal joint, and (2) the electrode providing relatively lower intensity stimulation was placed one to two cm below the fifth carpometacarpal joint. To determine individual thresholds for high-intensity and low-intensity stimulation, we followed a standard procedure (Brooks et al., 2010; Singer et al., 2006) and employed a visual analog rating scale (VAS) with endpoints defined as 0 = 'cannot feel anything' and 10 = 'maximum tolerable pain'. Tactile stimulation was delivered via two Digitimer DS5 isolated bipolar constant current stimulators (bipolar constant current, 5V, 50mA, Digitimer Ltd, Welwyn Garden City, UK) and a custom-made fMRI compatible 5-mm ring electrode, which delivered a maximally focused and centered tactile stimulus. By varying current

amplitude between 1 and 99 % of maximum amperage, stimuli with varying intensity levels were repeatedly delivered to each participant until stable ratings were achieved at least three times according to the following criteria: between 1 and 2 for the low intensity stimulus, and between 8 and 9 for the high intensity stimulus. Visual and tactile stimulus presentations, as well as recording of responses, were controlled by Cogent2000 (http://www.vislab.ucl.ac.uk/cogent.php).

Task

To investigate the effect of incidental affect on trust-taking, we employed a hybrid fMRI design, in which aversive affect was manipulated in a blocked fashion while social (trust) and non-social (control) tasks were presented in an event-related fashion. Specifically, we varied aversive affect by creating an expectancy of weak or strong unpredictable electrical stimulation that could occur at any time for the duration of an entire block. This expectancy was created by means of a block cue presented at the beginning of each block that informed participants about the game type (trust or control game) and the intensity of stimulation (weak or strong) for the current block (Figure C-1). Stimulation intensity was communicated to subjects in three ways: (1) via a verbal cue embedded in the 750-ms block cue [“strong” for treatment (“threat” condition), “weak” for control (“no-threat” condition)]; (2) via a tactile reminder cue presented 700 ms after visual cue onset that reflected the exact stimulation intensity of the current block; (3) via a specific background color that was consistently associated with either threat or no-threat blocks for each subject (color was counterbalanced across subjects) and remained constant for the duration of a block. The number and time points of electrical stimulation events throughout the blocks were determined to be completely unpredictable to subjects, in order to augment the efficacy of the threat-of-shock treatment. For this purpose, the number of stimulation events was determined for each block by random draw from a gamma distribution (shape parameter = 1; scale parameter = 1). The exact timing of these stimulation events was then determined at random time points between the offset of the cue display and onset of the resting screen drawing from a uniform distribution, with the constraint that at least 0.2 s separated successive electrical shocks.

Each block commenced with the set of cues described above that indicated the type of decision to be made (non-social control or trust) and the level of stimulation (weak, strong) to be expected by subjects for the rest of the block. After a brief and jittered interstimulus interval of 3-9 seconds, the first trial was displayed. In both the trust and the

control game, subjects were presented with a multiple-choice scenario, in which one of five amounts between 0 and 24 Swiss Francs (CHF) could be transferred to Player B or invested in a lottery. While subjects always had the options to either invest all (24 CHF) or none (0 CHF) of their endowment, each trial presented a novel choice scenario by (1) varying the intermediate options between 4, 6, or 8 CHF in the low category, 10, 12, or 14 CHF in the medium category, and 16, 18, or 20 in the high category of intermediate transfer amounts; (2) varying the location of each choice option and (3) varying the location of the originally highlighted choice option. This variability was introduced in order to ensure that subjects paid attention to all choice options on every trial and to avoid excessive use of heuristics. Intermediate amounts, location of choice options and location of the initially highlighted choice option were fully counterbalanced across conditions. Subjects selected their preferred option by moving a yellow dot that highlighted the currently selected choice option up and down by pressing two dedicated buttons on a standard MR-compatible 4-button response box and confirming their choice by pressing a third button. At this point the selected choice option was highlighted in red for the remaining duration of a trial. After a jittered intertrial interval (3-9 seconds) a new trial began.

Payment determination

We collected trustee responses in separate behavioral sessions that were conducted prior to the fMRI experiment using the same trust game. We elicited the trustees' choices with the strategy method, i.e. the trustees indicated their responses to each feasible transfer level. The trustees gave written and informed consent that we could use their strategies in follow-up experiments. In the fMRI part of our experiment the subjects (investors) thus played against the pre-recorded strategies of the trustees, i.e. a subject's transfer level together with the strategy of the (randomly) matched trustee determined the final monetary outcome in a trust game trial. Given the absence of the trustee on the scanning day, we informed participants that they were interacting with trustees in a temporally delayed fashion. Specifically, we emphasized to subjects that their payoffs were determined by decisions of real persons in the trust game, and by a computer algorithm in the control game, and that they were assigned different real persons across trust game trials. Finally, to maintain the interactive nature of the trust game, we informed our subjects that their choices had real, but delayed, consequences for trustees, who were sent additional payments according to the decisions made by the investor in the scanner after completion of the experiment. During

the experiment, the subjects did not receive any feedback about the behavior of their matched trustees, or the payoff amounts from lottery investments.

After completion of the fMRI part of the experiment, the subjects selected two trials at random by dice throw and payment was determined according to the decisions made by the subject and the trustee on the selected trials. In order to avoid hedging, both payout trials were drawn from the entire experiment, i.e. the payout trials were not specific to a condition, such as the trust or control game. If a trust game was randomly chosen for payout determination, the investor's payout was determined based on the amount transferred to the trustee and the backtransfer amount of the specific trustee the investor was paired with on that trial (payout investor = 24 - transfer to trustee + backtransfer from trustee; payout trustee = 24 + transfer from investor * 3 - backtransfer to investor). If a control game was randomly chosen, the computer algorithm randomly drew a payout amount from the distribution of trustees' backtransfer amounts. Our procedure therefore created equivalent payout amounts and likelihoods for the trust and control game.

Exit questionnaire

After completion of the experiment, subjects filled out an exit questionnaire that probed their beliefs about the accuracy of our instructions, as well as emotional reactions to our experimental manipulations. The main goal of the exit questionnaire was to measure whether subjects believed our instructions. Note that we implement such measurements routinely although we have little reason to believe that subjects doubt our instructions. Our laboratory uses deception only as a very rare exception, and we also did not use any deception in this experiment and fully disclosed all information truthfully to the subjects. Subjects were asked to rate 7 statements on a scale from 0, indicating very unbelievable, to 4, indicating very believable. The statements declared that the trust games were played with real persons, that each trust game was played with an anonymous trustee, that decisions of trustees were made by actual persons, and that trustees will receive additional payments based on the decisions of subjects on the relevant trial. Subjects' responses were entered into one-sample t-tests testing whether responses were significantly greater than the mid-point of the scale (2, indicating neither believable nor unbelievable). Mean ratings for all statements were significantly greater than two, indicating that subjects believed the statements (all t-tests survive the Bonferroni-corrected threshold of 0.007; average rating (SD) over all statements is 3.37 (0.86).

Skin conductance responses (SCRs)

Skin conductance responses were collected using a PowerLab 4/25T amplifier with a GSR Amp (ML116) unit and a pair of MR-compatible finger electrodes (MLT117F), which were attached to the participants' left middle and ring finger via dedicated Velcro straps after application of conductance gel. Subjects' hands had been washed using soap without detergents before the experiment. Stable recordings were ensured before starting the main experiment by waiting for signal stabilization during training and stimulation intensity calibration. LabChart (v. 5.5) software was used for recordings, with the recording range set to 40 μS and using initial baseline correction ("subject zeroing") to subtract the participant's absolute level of electrodermal activity from all recordings (all specs for devices and electrodes from ADInstruments Inc., Sydney, Australia).

Due to technical problems, data from 4 subjects included 1 run (out of 2) and data from 1 subject was lost. Each participant's SCR data were initially smoothed with a running average over 500 samples (equivalent to 500 ms at a sampling rate of 1KHz) to reduce scanner-induced noise. Data were then resampled from 1KHz to 1Hz and subsequently z-transformed. Statistical analysis of the pre-processed skin conductance data followed the approach commonly employed in analyses of fMRI data. Specifically, multiple linear regression implemented in AFNI was used to estimate SCR during decisions made in each of the task conditions, that is, during trust and non-social control tasks and in the context of threat and no-threat treatment blocks. The statistical model included a total of 7 regressors that reflected the onset times of decision screens in trust and non-social control trials under expectancy of strong and weak electrical shocks, cue times indicating the onset of a block, as well as delivery times of strong and weak tactile stimulation. To avoid making assumptions about the shape of the SCR response, a finite impulse response (FIR) model was used to estimate average responses (beta weights) during each trial type via deconvolution from event onset to 16s post onset using 17 cubic spline basis functions. Constant, linear and quadratic terms were included as regressors of no interest for each run separately to model baseline drifts of the SCR. Regressor estimates (beta weights) at each time point and for each condition were then used in follow-up analyses reported in the Results section.

fMRI data acquisition

Magnetic resonance images were collected using a 3T Philips Intera whole-body magnetic resonance scanner (Philips Medical Systems, Best, The Netherlands) equipped with an 8-

channel Philips sensitivity-encoded (SENSE) head coil. Structural image acquisition consisted of 180 T1-weighted transversal images (0.75-mm slice thickness). For functional imaging, a total of 1095 volumes were obtained using a SENSE T2*-weighted echo-planar imaging sequence (Pruessmann et al., 1999) with an acceleration factor of 2.0. We acquired 45 axial slices covering the whole brain with a slice thickness of 2.8 mm (inter-slice gap of 0.8 mm, sequential acquisition, repetition time = 2470 ms, echo time = 30 ms, flip angle = 82°, field of view = 192 mm, matrix size = 68 × 68). To optimize functional sensitivity in orbitofrontal cortex and medial temporal lobes, we used a tilted acquisition in an oblique orientation at 15° relative to the AC-PC line.

fMRI data analysis

Preprocessing and statistical analyses were performed using SPM8 (Wellcome Department of Imaging Neuroscience, London, UK). To correct for head motion, all functional volumes were realigned to the first volume using septic b-spline interpolation and subsequently unwarped to remove residual movement-related variance due to susceptibility-by-movement interactions. Slice timing correction was performed after realignment/unwarping. To improve coregistration, bias-corrected anatomical and mean EPI images were created and subsequently coregistered using the new segment toolbox in SPM. Images were normalized to the Montreal Neurological Institute T1 template using the parameters (forward deformation fields) derived from the nonlinear normalization of individual gray matter tissue probability maps. Finally, functional data underwent spatial smoothing using an isotropic 6 mm FWHM Gaussian kernel.

Statistical analyses were carried out using the general linear model. Regressors of interest were modeled using a canonical hemodynamic response function (HRF) with time and dispersion derivatives in order to account for subject-to-subject and voxel-to-voxel variation in response peak and dispersion (Henson et al., 2002). Since our main interest was the impact of aversive affect on trust-taking, we modeled the decision period for the full response time on each trial, that is from the onset of the decision screen until subjects pressed the confirm button. This was done in the following four conditions: (1) trust game during relatively high-intensity stimulation expectancy (threat condition), (2) trust game during relatively low-intensity stimulation expectancy (no-threat condition), (3) control game during relatively high-intensity stimulation expectancy (threat condition) and (4) control game during relatively low-intensity stimulation expectancy (no-threat condition). Finally, the following regressors of no interest were included in our model: the actually

realized weak and strong tactile stimulations during a block (one reminder shock and, on average, one additional shock randomly drawn from a gamma distribution were administered per block), block cues indicating game type (trust, control) and stimulation intensity of the reminder shock (weak, strong) at the beginning of each block, as well as omissions of behavioral responses during a trial.

The main goal of the current investigation was to identify the impact of aversive affect on the neural correlates of trust decisions. Trust-specific neural effects of aversive affect can be identified via an interaction between threat and choice domain, in which threat significantly alters the neural correlates of decision-making in trust relative to non-social control trials. To investigate the interaction between threat and choice domain, an ANOVA was computed by entering contrast estimates obtained from first level models into a flexible factorial model with the factors choice domain (trust, control), threat (absent, present), as well as subject. We were particularly interested in trust-specific affect-induced suppression of activity and connectivity, which we tested via the interaction contrast ($Trust_{No\ threat}$ > $Trust_{Threat}$) > (NS $Control_{No\ threat}$ > NS $Control_{Threat}$) in the context of the flexible factorial design. A covariate reflective of each subject's mean transfer in each condition was also included to probe for brain-behavior correlations. All analyses were also conducted without the behavioral covariate and results did not change.

We expected regions commonly implicated in the major cognitive and affective component processes of trust taking to be affected by aversive affect. ROI analyses in relevant cortical regions were conducted using small volume correction with masks created via relevant search terms on neurosynth.org (Yarkoni et al., 2011). Neurosynth.org offers a means to obtain automated meta-analyses over a large number of prior fMRI investigations and thereby provides an independent method to obtain masks for ROI analyses. Our hypothesized component processes implicate the following regions in processing specific aspects of decision-making: 1. Evaluating the anticipated outcomes of choice options, which involves regions commonly implicated in reward processing and valuation, such as ventral striatum and vmPFC (Levy and Glimcher, 2012) (neurosynth search: "reward"), 2. Assessing the trustworthiness of the trustee to make predictions about payout probability in the trust game, which involves regions commonly implicated in theory of mind and social cognition, such as TPJ and dmPFC (Van Overwalle, 2009) (neurosynth search: "theory of mind"), and 3. Regions implicated in emotion processing, such as the amygdala (Lindquist et al., 2012) (neurosynth search: "emotion"). We entered the search terms specified above into neurosynth to extract cortical ROI masks. ROI analyses in anatomically well-defined

subcortical regions, such as the amygdala, as well as the caudate head and putamen in ventral striatum were conducted using anatomical masks created via the AAL atlas implemented in WFU Pickatlas. Taken together, we employed the following masks in region of interest analyses: (1) bilateral temporoparietal junction (neurosynth: theory of mind), with peak voxels in left (-60,-56,14) and right TPJ (56,-58,20) and sizes of 1031 and 1416 voxels, respectively; (2) bilateral amygdala (AAL) with sizes of 439 (left) and 492 (right) voxels; (3) bilateral ventral striatum (combined mask of AAL putamen and caudate up to z = 8), with sizes of 3239 (left) and 3429 (right) voxels, respectively; (4) dorsomedial PFC (neurosynth search term: theory of mind), with a peak voxel in medial dmPFC (-2, 28, 62) and a size of 3175 voxels; (5) ventromedial PFC (neurosynth search term: ventromedial) with a peak voxel in medial vmPFC (8, 24, -12), with a size of 1327 voxels. For the analysis of trust decisions, we employed a combination of valuation- (vmPFC, vStr), social cognition- (TPJ, dmPFC) and emotion- (amygdala) related ROIs.

Furthermore, to identify whether extended networks outside our regions of interest show effects of interest, we conducted whole brain analyses at an FWE-corrected extent threshold of $p < 0.05$ ($k > 499$, initial cluster-forming height threshold $p < 0.005$). Finally, to characterize activation patterns of interest, such as time courses and activation differences due to aversive affect, regression coefficients (beta weights) for the canonical HRF regressors were extracted with rfxplot (Gläscher, 2009) from a 6 mm sphere around the peak voxel that showed significant effects of interest on BOLD responses and functional connectivity. Follow-up tests that characterize the single components of significant interaction effects were conducted in neuroimaging space via tests of simple effects of interest.

PPI analyses

Psychophysiological Interaction analyses were conducted using the generalized form of context-dependent psychophysiological interactions toolbox (gPPI toolbox; McLaren et al., 2012), using the same statistical model as outlined above. All voxels that survived SV FWE-correction for the interaction contrast in left TPJ (-60, -54, 19, k=95) were used as seed region (shown in blue color in Figure C-3c). To obtain an estimate of neural activity within the seed region, the BOLD signal from the seed region was extracted, the signal was corrected by removing effects of noise covariates and the corrected signal was deconvolved (Gitelman et al., 2003). Psychological interaction regressors for each of the task type and stimulation intensity combinations [control decisions during (1) weak and (2) strong

stimulation, trust decisions during (3) weak and (4) strong stimulation] were created by multiplying the estimated neural activity during the relevant decisions with condition-specific on- and offset times convolved with the canonical HRF. A new GLM was then estimated for each subject that consisted of the original design matrix with the addition of the four psychological interaction regressors and the time course from the seed region. To investigate the impact of aversive affect on trust-specific functional connectivity with relevant seed regions, we probed the functional connectivity data for an interaction between threat and choice domain. To investigate the interaction between threat and choice domain, an ANOVA was computed by entering contrast estimates obtained from first level models into a flexible factorial model with the factors choice domain (trust, non-social control), threat (absent, present), as well as subject. To investigate trust-specific brain-behavior correlations due to threat, we estimated two additional flexible factorial models with (a) the factor choice domain and (b) the factor threat and (c) importantly, separate covariates reflecting mean transfer in each condition that were interacted with the factor of interest (choice domain, or threat). A subject factor was included in all models. Given that we were particularly interested in trust-specific changes in functional connectivity, we first contrasted the covariate reflecting mean trust to that for mean transfer in the non-social control task. In the context of our model, a significant result for a region indicates a *difference* between the trust and the control task in the (linear) relationship between the functional connectivity with the TPJ and mean transfer levels. In Figure C-4 we illustrate these results by extracting regression coefficients reflecting functional connectivity strength in each of the conditions from 6-mm spheres around interaction peak voxels and performing the following equivalent regression analysis:

$$y_{ik} = \beta_0 + \beta_1\, Transfer_{ik} + \beta_2\, Choice\ Domain_k + \beta_3\, (Transfer_{ik} * Choice\ Domain_k) + \varepsilon_{ik}$$

where the dependent variable y_{ik} is the functional connectivity strength observed within a given brain region for individual *i* in Choice Domain k. $Transfer_{ik}$ is the mean amount sent by individual *i* in Choice Domain k and Choice Domain is a dummy variable encoding whether decisions were made in the trust or the control task (1 indicates trust, 0 indicates control task). In this regression, the coefficient for $Transfer_{ik}$ measures the slope of the relationship between TPJ connectivity and mean transfers in the control task (see blue lines in Figures C-4a – c), and the sum of the coefficients for $Transfer_{ik}$ and the interaction term ($Transfer_{ik}$ * $Choice\ Domain_k$) measures the slope of the relationship between TPJ

connectivity and mean transfers in the trust task (see orange lines in Figures C-4a – c). Equivalent analyses were performed to probe for significant differences between the threat and the no-threat condition in the trust game in the (linear) relationship between functional TPJ connectivity and mean transfer levels.

C.6 References

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., and Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. Neuron *58*, 639–650.

Behrens, T.E.J., Hunt, L.T., and Rushworth, M.F.S. (2009). The computation of social behavior. Science *324*, 1160–1164.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, Reciprocity, and Social History. Games Econ. Behav. *10*, 122–142.

Blackburn, I.M., Cameron, C.M., and Deary, I.J. (1990). Individual differences and response to the Velten mood induction procedure. Personal. Individ. Differ. *11*, 725–731.

Bogdan, R., and Pizzagalli, D.A. (2006). Acute Stress Reduces Reward Responsiveness: Implications for Depression. Biol. Psychiatry *60*, 1147–1154.

Bohnet, I., and Zeckhauser, R. (2004). Trust, risk and betrayal. Trust Trust. *55*, 467–484.

Brooks, A.M., Pammi, V.S.C., Noussair, C., and Berns, G.S. (2010). From bad to worse: striatal coding of the relative value of painful decisions. Front. Decis. Neurosci. *4*, 176.

Buchwald, A.M., Strack, S., and Coyne, J.C. (1981). Demand characteristics and the Velten mood induction procedure. J. Consult. Clin. Psychol. *49*, 478–479.

Delgado, M.R., Frank, R.H., and Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. Nat. Neurosci. *8*, 1611–1618.

Fehr, E. (2009). On the Economics and Biology of Trust. J. Eur. Econ. Assoc. *7*, 235–266.

Frost, R.O., and Green, M.L. (1982). Velten Mood Induction Procedure Effects Duration and Postexperimental Removal. Pers. Soc. Psychol. Bull. *8*, 341–347.

Gitelman, D.R., Penny, W.D., Ashburner, J., and Friston, K.J. (2003). Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. NeuroImage *19*, 200–207.

Gläscher, J. (2009). Visualization of group inference data in functional neuroimaging. Neuroinformatics *7*, 73–82.

Grabenhorst, F., and Rolls, E.T. (2011). Value, pleasure and choice in the ventral prefrontal cortex. Trends Cogn. Sci. *15*, 56–67.

Grillon, C., Ameli, R., Foot, M., and Davis, M. (1993). Fear-potentiated startle: relationship to the level of state/trait anxiety in healthy subjects. Biol. Psychiatry *33*, 566–574.

Gromann, P.M., Heslenfeld, D.J., Fett, A.-K., Joyce, D.W., Shergill, S.S., and Krabbendam, L. (2013). Trust versus paranoia: abnormal response to social reward in psychotic illness. Brain J. Neurol. *136*, 1968–1975.

Hariri, A.R., Mattay, V.S., Tessitore, A., Fera, F., and Weinberger, D.R. (2003). Neocortical modulation of the amygdala response to fearful stimuli. Biol. Psychiatry *53*, 494–501.

Harlé, K.M., Chang, L.J., van 't Wout, M., and Sanfey, A.G. (2012). The neural mechanisms of affect infusion in social economic decision-making: a mediating role of the anterior insula. NeuroImage *61*, 32–40.

Henson, R.N.A., Price, C.J., Rugg, M.D., Turner, R., and Friston, K.J. (2002). Detecting Latency Differences in Event-Related BOLD Responses: Application to Words versus Nonwords and Initial versus Repeated Face Presentations. NeuroImage *15*, 83–97.

Isen, A.M., Clark, M., and Schwartz, M.F. (1976). Duration of the effect of good mood on helping: "Footprints on the sands of time." J. Pers. Soc. Psychol. *34*, 385–393.

Kable, J.W., and Glimcher, P.W. (2009). The neurobiology of decision: consensus and controversy. Neuron *63*, 733–745.

King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., and Montague, P.R. (2008). The rupture and repair of cooperation in borderline personality disorder. Science *321*, 806–810.

Knutson, B., Wimmer, G.E., Kuhnen, C.M., and Winkielman, P. (2008). Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. Neuroreport *19*, 509–513.

Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., and Fehr, E. (2005). Oxytocin increases trust in humans. Nature *435*, 673–676.

Kuhnen, C.M., and Knutson, B. (2011). The Influence of Affect on Beliefs, Preferences, and Financial Decisions. J. Financ. Quant. Anal. *46*, 605–626.

LeDoux, J. (2007). The amygdala. Curr. Biol. CB *17*, R868–R874.

Lerner, J.S., Li, Y., and Weber, E.U. (2013). The financial costs of sadness. Psychol. Sci. *24*, 72–79.

Levy, D.J., and Glimcher, P.W. (2012). The root of all value: a neural common currency for choice. Curr. Opin. Neurobiol. *22*, 1027–1038.

Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., and Barrett, L.F. (2012). The brain basis of emotion: a meta-analytic review. Behav. Brain Sci. *35*, 121–143.

Loewenstein, G. (2000). Emotions in Economic Theory and Economic Behavior. Am. Econ. Rev. *90*, 426–432.

Marston, A., Hart, J., Hileman, C., and Faunce, W. (1984). Toward the Laboratory Study of Sadness and Crying. Am. J. Psychol. *97*, 127–131.

Martin, M. (1990). On the induction of mood. Clin. Psychol. Rev. *10*, 669–697.

McLaren, D.G., Ries, M.L., Xu, G., and Johnson, S.C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. NeuroImage *61*, 1277–1286.

Mitchell, J.P., Macrae, C.N., and Banaji, M.R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. Neuron *50*, 655–663.

Morishima, Y., Schunk, D., Bruhin, A., Ruff, C.C., and Fehr, E. (2012). Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. Neuron *75*, 73–79.

Ochsner, K.N., Ray, R.D., Cooper, J.C., Robertson, E.R., Chopra, S., Gabrieli, J.D.E., and Gross, J.J. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. NeuroImage *23*, 483–499.

Ochsner, K.N., Silvers, J.A., and Buhle, J.T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. Ann. N. Y. Acad. Sci. *1251*, E1–E24.

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. Hum. Brain Mapp. *30*, 829–858.

Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a "low road" to "many roads" of evaluating biological significance. Nat. Rev. Neurosci. *11*, 773–783.

Phelps, E.A. (2009). The study of emotion in neuroeconomics. In Neuroeconomics: Decision Making and the Brain, P.W. Glimcher, C. Camerer, E. Fehr, and R.A. Poldrack, eds. (New York: Academic Press),.

Polivy, J., and Doyle, C. (1980). Laboratory induction of mood states through the reading of self-referent mood statements: Affective changes or demand characteristics? J. Abnorm. Psychol. *89*, 286–290.

Pruessmann, K.P., Weiger, M., Scheidegger, M.B., and Boesiger, P. (1999). SENSE: sensitivity encoding for fast MRI. Magn. Reson. Med. Off. J. Soc. Magn. Reson. Med. Soc. Magn. Reson. Med. *42*, 952–962.

Rangel, A., and Hare, T. (2010). Neural computations associated with goal-directed choice. Curr. Opin. Neurobiol. *20*, 262–270.

Saxe, R., and Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind." NeuroImage *19*, 1835–1842.

Schmitz, A., and Grillon, C. (2012). Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test). Nat. Protoc. *7*, 527–532.

Schulreich, S., Heussen, Y.G., Gerhardt, H., Mohr, P.N.C., Binkofski, F.C., Koelsch, S., and Heekeren, H.R. (2014). Music-evoked incidental happiness modulates probability weighting during risky lottery choices. Front. Psychol. *4*, 981.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.

Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, A.S., Winter, J.J., and Davidson, R.J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. Nat. Rev. Neurosci. *12*, 154–167.

Singer, T., Seymour, B., O'Doherty, J.P., Stephan, K.E., Dolan, R.J., and Frith, C.D. (2006). Empathic neural responses are modulated by the perceived fairness of others. Nature *439*, 466–469.

Wager, T.D., Davidson, M.L., Hughes, B.L., Lindquist, M.A., and Ochsner, K.N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. Neuron *59*, 1037–1050.

Westermann, R., Spies, K., Stahl, G., and Hesse, F.W. (1996). Relative effectiveness and validity of mood induction procedures: a meta-analysis. Eur. J. Soc. Psychol. *26*, 557–580.

Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. Neuron *40*, 655–664.

Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., and Wager, T.D. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nat. Methods *8*, 665–670.

Young, L., Camprodon, J.A., Hauser, M., Pascual-Leone, A., and Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. Proc. Natl. Acad. Sci. 200914826.

C.7 Supplemental information

Supplementary Figures 1-4

a b

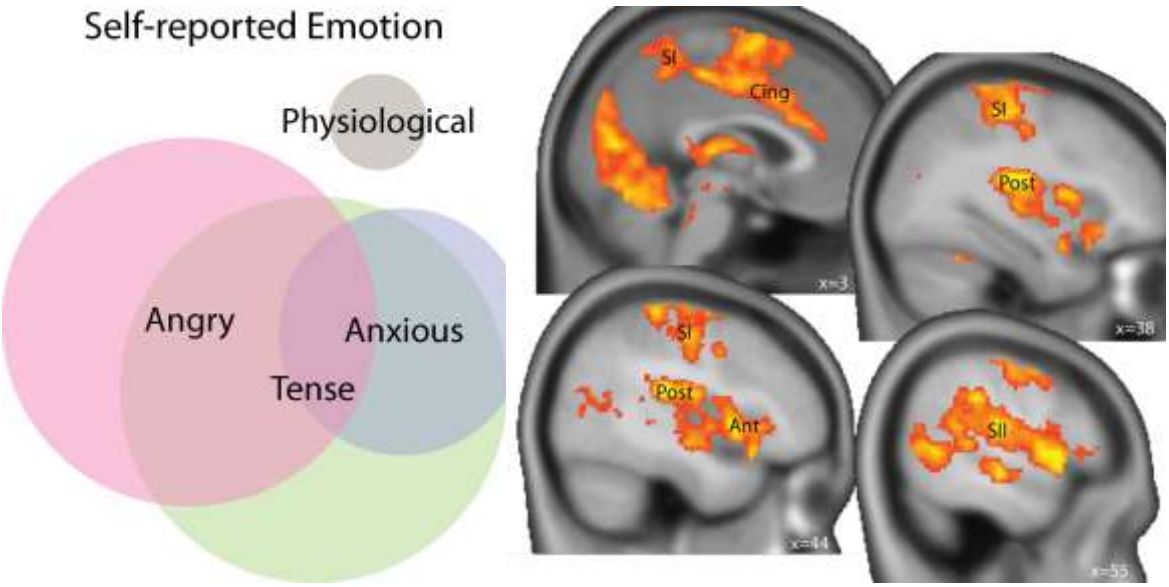

Figure S1: Manipulation checks (related to Figure C-2). (a) Self-reported experience of emotion during threat blocks. Venn diagram illustrating proportions and overlap between self-reported emotional reactions to painful tactile stimulation, in which subjects were able to report multiple emotions. We asked subjects to report how they felt during threat blocks in an open-ended exit questionnaire. 95.12% of subjects indicated that they experienced aversive emotional arousal during pain blocks, such as tenseness and/or anger and/or sadness. To illustrate the frequency of emotional reactions to painful tactile stimulation, subjects' answers were binned into three emotion categories that best summarize their emotional responses: angry and annoyed responses were grouped into the "Angry" category (25), tense,

stressed and surprised responses were grouped into the "Tense" category (27), and scared, nervous, helpless, and sad responses were grouped into the "Anxious" category (11). Most subjects reported angry (61%) and tense (66%) emotions while a minority also felt anxious (27%). Of note, due to the open-ended nature of the questions, subjects were able to indicate more than one emotional reaction, as illustrated by the overlap between different sets in the Venn diagram. (b) Administration of painful compared to just-noticeable tactile stimulation led to increased activity within key regions of the pain matrix, including primary (SI) and secondary (SII) somatosensory cortex, anterior (Ant) and posterior (Post) insula, as well as mid cingulate cortex (Cing). The contrast images reflect activation at the time of strong vs. weak tactile stimulation and are thresholded at $p < 0.05$ FWE-corrected at the cluster level (with a cluster-forming voxel-level threshold of $p < 0.001$).

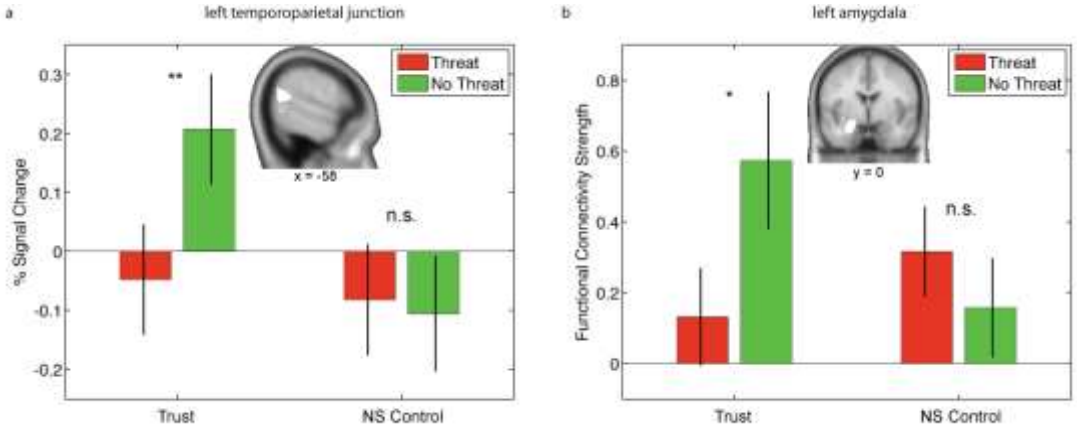

Figure S2: Additional post-hoc inspection of the significant interactions reported in the main paper within all voxels of independent TPJ and amygdala masks (related to Figure C-3). To characterize the interaction patterns reported in the main paper, we extracted subject-specific regression coefficients (beta weights) from all voxels within the entire left TPJ (k = 1031) and amygdala (k = 439) masks and performed post-hoc statistical analyses in independent ROIs. (a) Post-hoc pairwise comparisons in the TPJ ROI revealed significantly greater choice-related activation during no-threat relative to threat when subjects made decisions in the trust task [$t(40) = 2.769$, $p = 0.0085$], but not when they made decisions in the control task [$t(40) = -0.167$, $p = 0.8682$]. (b) Post-hoc pairwise comparisons in the amygdala ROI revealed significantly greater choice-related activation during no-threat relative to threat when subjects made decisions in the trust task [$t(40) = 2.097$, $p = 0.0424$], but not when they made decisions in the control task [$t(40) = -0.812$, $p = 0.4216$]. These results confirm that the presence of aversive affect during trust decisions led to a significant suppression of (a) TPJ activity and (b) TPJ-amygdala connectivity relative to the presence of aversive affect. ** $p < 0.01$; * $p < 0.05$; n.s.: not significant

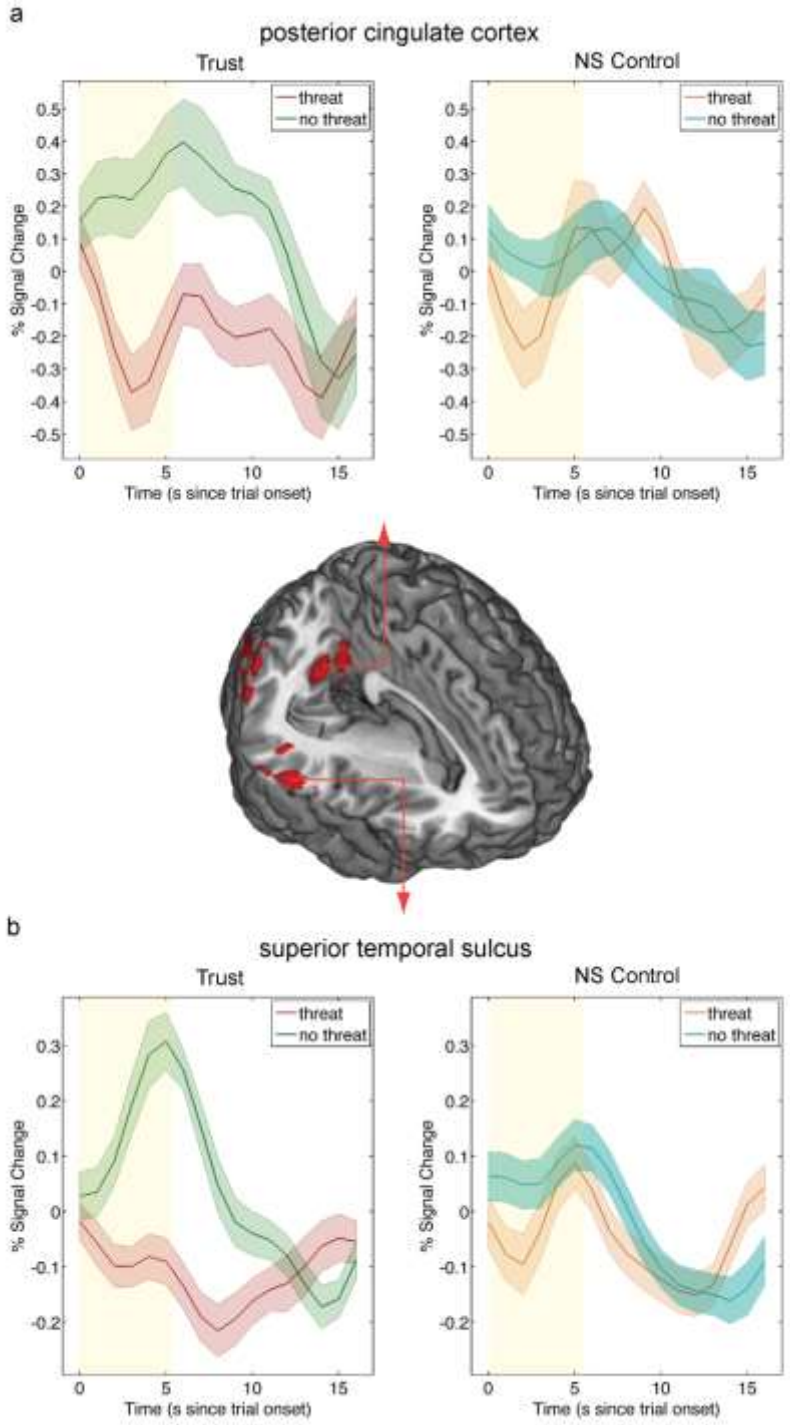

Figure S3: The impact of aversive affect during trust trials in those regions that remain after we remove regions that show a domain-general effect of threat (related to Figure C-3). Regions that showed a domain general effect were removed via an exclusive mask identified by the main effect of threat. ROI analyses (Supplementary Table 4) show that aversive affect causes a significant deactivation in vmPFC (8, 30, -11, k = 72), and ventral striatum (-3, 3, -6, k = 40). Whole-brain analysis at cluster-level FWE-correction investigated whether an *extended network outside our a-priori regions of interest* shows a specific impact of aversive affect on the neural correlates of trust decisions. This additional analysis (Supplementary Table 5) revealed that the threat condition caused a reduction of activations relative to no-threat in (a) posterior cingulate cortex [-6, -46, 7, k = 859] and (b) left superior temporal sulcus [-63, -25, 4, k = 1322], and in left inferior parietal lobe [-52, -60, 45, k = 691] which is not shown in Supplementary

Figure 4. Time courses in (a) and (b) illustrate a significant suppression due to aversive affect during trust decisions (green > red), but not during the control task (aquatic green = orange).

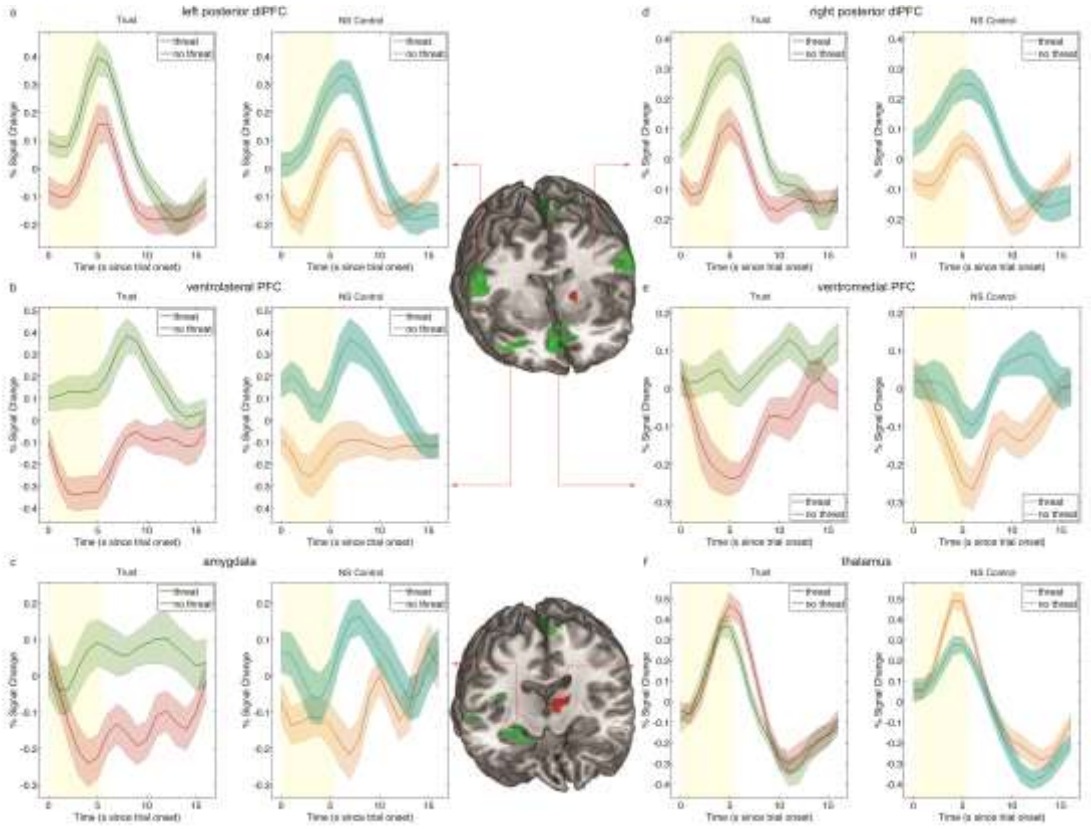

Figure S4: The impact of aversive affect on choice-domain independent neural correlates of decision-making. A domain-general network consisting of bilateral posterior dlPFC [(a) left: -62, -4, 18, k = 1901 and (d) right: 62, -6, 28, k = 1010], left amygdala [(c) -24, -15, -23, k = 552], posterior paracentral lobule [not shown, 4 -36, 69, k = 887], and a large cluster in ventral anterior prefrontal cortex (k = 4082) that includes left vlPFC [(b) -48, 41, -8] and vmPFC [(e) -10, 44, -8], shows significant threat-related reduction (no threat > threat, regions shown in green) in choice-related activity during both trust and non-social control trials. Significant threat-related enhancement of activity (threat > no threat, regions shown in red) during the decision phase was obtained in the thalamus [(f), 18, -6, 1, k = 559] and cerebellum [not shown, -4, -46, -24, k = 849]. Time courses illustrate similar suppressions (a – e) and enhancements (f) due to the aversive affect during decisions in trust and non-social control trials. Time courses are shown for both trust (left, threat shown in red, no threat shown in green) and control (right, threat shown in orange, no threat shown in aqua green) trials in separate graphs and were extracted from 6 mm spheres around peak voxels. The 5.5-second choice period is displayed in yellow.

Supplemental Tables 1-7

Table S1: OLS Regression results reflecting the influence of experienced electrical stimulation on choice behavior (related to Figure C-2d, see text S1 for more information on statistics employed)

Dependent Variable: Amount Transferred						
	(1)		(2)		(3)	
Threat	-0.821	****	-0.928	****	-0.866	****
	(0.238)		(0.255)		(0.217)	
Unpredictable S	0.627		0.640			
	(0.519)		(0.507)			
Threat x Unpredictable S	-0.528		-0.550			
	(0.710)		(0.714)			
Choice Domain	2.825	****	2.827	****	2.715	****
	(0.589)		(0.590)		(0.589)	
Choice Domain x Unpredictable S	-0.935		-0.879			
	(0.702)		(0.661)			
Threat x Choice Domain x Unpredictable S	0.536		0.424			
	(0.945)		(0.917)			
Predictable S	0.243				0.359	
	(0.489)				(0.486)	
Threat x Predictable S	-0.488				-0.614	
	(0.687)				(0.687)	
Choice Domain x Predictable S	0.256				0.084	
	(0.588)				(0.570)	
Threat x Choice Domain x Predictable S	-0.498				-0.325	
	(0.841)				(0.813)	
Gender	1.265		1.264		1.266	
	(1.372)		(1.372)		(1.372)	
Age	0.264		0.264		0.264	
	(0.330)		(0.330)		(0.330)	
(Intercept)	12.244	****	12.298	****	12.326	****
	(0.911)		(0.915)		(0.928)	
F	18.67		27.82		27.8	
P	<0.001		<0.001		<0.001	
N	3437		3437		3437	
Corrected AIC	22717		22711		22711	

This table reports OLS coefficient estimates (robust standard errors corrected for clustering on the individual level in parentheses). The dependent variable in both columns is the amount invested in either Player B or the lottery. Column (1) contains the results from the full regression that includes both predictable shocks (i.e., reminder shocks) and unpredictable shocks and all interactions, while columns (2) and (3) reflect results from reduced models that exclude the influence of predictable (2) and unpredictable (3) shocks. "Threat" is a dummy variable reflecting the *expectation* of impending painful electric shock during the current choice scenario. "Choice Domain" is a dummy variable that indicates

trust trials. "Unpredictable S" and "Predictable S" are dummy variables encoding the experience of at least one unpredictable and, respectively, predictable shock in the interval from 5 seconds prior to the display of the choice screen until button press. We also investigated shorter and longer intervals from 1 to 10 seconds prior to the choice scenario in 1-second steps and in all cases only Threat and Choice Domain reach statistical significance. "Age" is the mean-centered individual's age in years, and "Gender" is a gender dummy encoding the influence of being male on transfer rates. The number of observations is less than the total number of choice scenarios (3444) due to omissions (response times > 5.5 seconds). Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, **** $p < 0.001$

Table S2: Regions of interest analyses investigating trust-specific neural correlates in regions associated with social cognition and valuation (related to Figure C-3a and b). (a) Regions that are preferentially involved in the trust game relative to the non-social control task when aversive affect is absent. (b) Only the temporoparietal junction (TPJ) shows a significant interaction effect reflective of a change in the neural impact of aversive affect during trust taking relative to the non-social control task. (c) Regions that show significant simple effects of threat during trust decisions.

Structure	L/R	Cluster Size	Peak t	x	y	z	Peak p
(a) Simple effect: No threat (Trust > NS Control)							
TPJ	left	247	4.34	-57	-60	27	0.003
dorsomedial PFC	left	39	3.85	-9	60	18	0.044
ventromedial PFC	bil.	89	3.88	9	32	-12	0.031
(b) Interaction: Trust (No threat > Threat) – NS Control (No threat > Threat)							
TPJ	left	95	3.63	-60	-54	19	0.035
(c) Simple effect: Trust (No threat > Threat)							
TPJ	left	103	4.03	-58	-55	19	0.012
vmPFC	bil.	72	3.94	8	30	-11	0.026
ventral striatum	left	23	3.76	-26	9	-9	0.048

Activation clusters survive small volume correction for multiple comparisons based on FWE correction at the peak level for a priori regions of interest using masks created with reverse inference maps from relevant search terms on the neurosynth.org database. The left TPJ masks contained 1031 voxels; the dorsomedial and ventromedial PFC masks contained 3175 and 1327 voxels, respectively, and the left ventral striatum mask contained 3239 voxels.

Table S3: Whole brain analysis investigating the neural effects of aversive affect (related to Figure S3). Results from whole brain analysis investigating the neural effects of aversive affect (i.e., no threat vs. threat condition) for trust decisions after removing regions that show a domain-general effect of threat (SE-ME, $p < 0.005$, cluster size > 499, FWE corrected at cluster-level).

Structure	L/R	Cluster Size	Peak t	x	y	z
inferior parietal cortex	left	691	4.3	-52	-60	45
superior temporal sulcus	left	1322	4.19	-63	-25	4
posterior cingulate cortex	bil.	859	4.07	-6	-46	7

Table S4: Region of interest analyses investigating TPJ-amygdala connectivity patterns (related to Figures C-3c – d). (a) TPJ-amygdala connectivity during decision-making in the trust relative to the control task *when aversive affect is absent* and (b – c) the differential impact of aversive affect on TPJ-amygdala connectivity in the trust task relative to the control task.

Structure	**L/R**	**Cluster Size**	**Peak t**	**x**	**y**	**z**	**Peak p**
(a) Simple effect: No threat (Trust > NS Control)							
amygdala	left	14	3.26	-22	-9	-15	0.039
(b) Interaction: Trust (No threat > Threat) – NS Control (No threat > Threat)							
amygdala	left	29	3.68	-26	0	-23	0.012
(c) Simple effect: Trust (No threat > Threat)							
amygdala	left	110	4.03	-28	-6	-14	0.004

Activation clusters survive small volume correction for multiple comparisons based on FWE correction at the peak level for a priori regions of interest using anatomical masks created with reverse inference maps from relevant search terms on the neurosynth.org database. The left amygdala mask contained 439 voxels.

Table S5: Region of interest analyses investigating brain-behavior relationships in the absence of threat (related to Figure C-4). ROI analyses investigated whether, in the absence of threat, there is a difference between the trust and the control task in the correlation of mean transfers with functional TPJ connectivity and its target regions.

Structure	**L/R**	**Cluster Size**	**Peak t**	**x**	**y**	**z**	**Peak p**
Amygdala	right	48	3.53	27	2	-21	0.026

Activation clusters survive small volume correction for multiple comparisons based on FWE correction at the peak level for a priori regions of interest using masks created with reverse inference maps from relevant search terms on the neurosynth.org database. The right amygdala mask contained 492 voxels.

Table S6: Whole-brain analyses investigating brain-behavior relationships (related to Figure C-4). Results from whole-brain analysis investigating whether there is a difference between the trust and the control task in the correlation of mean transfers with functional TPJ connectivity with its target regions, in (a) the absence of threat and (b) the presence of threat ($p < 0.005$, cluster size > 499, FWE corrected at cluster-level).

Structure	L/R	Cluster Size	Peak t	x	y	z
(a) In the absence of threat						
Superior Temporal Sulcus	right	1076	5.49	63	-45	6
ventrolateral PFC / IFG	left	695	4.93	-50	23	-8
Paracentral Lobule / SFG	left	827	4.75	-12	-15	75
dorsomedial PFC	bil.	5069	4.48	-3	18	55
ventrolateral PFC / IFG	right	544	4.17	57	20	10
dorsolateral PFC	right	1464	4.06	50	12	37
intraparietal Sulcus	right	532	4	46	-58	45
dmPFC / SFG	right	546	3.93	29	50	39
(b) in the presence of threat						
Superior Temporal Sulcus	right	514	4.97	64	-43	4

Table S7: Whole-brain analyses investigating the main effect of threat (related to Figure S4). Results from whole-brain analysis investigating the main effect of threat on neural activity during decision-making independent of choice domain ($p < 0.005$, cluster size > 499, FWE corrected at cluster-level).

Structure	L/R	Cluster Size	Peak t	x	y	z
(a) *No threat > Threat*						
ventromedial/ventrolateral PFc	bil.	4082	5.91	-14	68	3
posterior dorsolateral PFC	right	1010	4.67	62	-6	28
posterior dorsolateral PFC	left	1901	4.49	-62	-4	18
posterior paracentral lobule	bil.	887	4.36	4	-36	69
(b) *Threat > No threat*						
cerebellum	bil.	849	-2.61	-4	-46	-24
thalamus	right	559	-2.61	6	-7	4

Supplemental Text

Text S1: Additional Behavioral Analyses and Results (related to Figure C-2 and Table S1). It is possible that the main driving force of the behavioral change that we report in the main paper is not due to the affective impact of our threat of shock manipulation, but to the actual experience of electrical stimulation immediately prior to decision-making. To assess this possibility, we investigated the effects of experienced electrical stimulation on choice behavior in the trust and the non-social control game. To this end, we ran comprehensive ordinary least-squares (OLS) regression analyses in R version 2.5.12. These analyses predicted for each individual *i* the observed choice $T_{i,t}$ on trial *t* with the following equation:

$$T_{i,t} = \beta_0 + \beta_1*\text{Threat}_{i,t} + \beta_2*\text{Choice Domain}_{i,t} + \beta_3* \text{US}_{i,t} + \beta_4*\text{PS}_{i,t} + \beta_5*\text{Threat}_{i,t}*\text{US}_{i,t} + \beta_6*\text{Choice Domain}_{i,t} \text{ x } \text{US}_{i,t} + \beta_7*\text{Threat}_{i,t} \text{ x Choice Domain}_{i,t} \text{ x } \text{US}_{i,t} + \beta_8*\text{Threat}_{i,t} \text{ x } \text{PS}_{i,t} + \beta_9*\text{Choice Domain}_{i,t} \text{ x } \text{PS}_{i,t} + \beta_{10}*\text{Threat}_{i,t} \text{ x Choice Domain}_{i,t} \text{ x } \text{PS}_{i,t} + \beta_{11}*X_i + \varepsilon_{i,t} \quad \text{(eq. 1)}$$

$T_{i,t}$ reflects the transfers to either a lottery or Player B on a given trial. $Threat_{i,t}$ is a dummy-coded variable (1 = threat of shock, 0 = safety) reflecting the *expectation* of impending painful electric shock during the current choice scenario, not the *experience* of shock. $Choice\ Domain_{i,t}$ is a dummy-coded variable (1 = trust game, 0 = control game) reflecting transfers within the context of a trust game. $US_{i,t}$ and $PS_{i,t}$ are dummy variables (1 = shock, 0 = no shock) encoding the *experience* of at least one unpredictable (US) and, respectively, predictable shock (PS) in the interval from 5 seconds prior to the display of the choice scenario until button press. We included PS in our regression models, in order to control for the influence of reminder shocks presented during the block cue. Of note, we also investigated shorter and longer intervals from 1 to 10 seconds prior to the choice scenario in 1-second increments and in all cases only *Threat* and *Choice Domain* reach statistical significance, no other factors or interactions. The model also contains a constant parameter (β_0), which measures the average transfer to the lottery in the no threat condition. Finally, the model includes a set of socio-economic control variables (X_i, e.g., age and gender) and relevant interaction terms that reflect differential effects of experiencing shocks on transfer behavior in different Choice Domain and emotion contexts. We employed a random-effects model with robust standard errors adjusted for clustering on the subject level.

The results show significant treatment (threat) and Choice Domain (trust) effects, which we also report in the ANOVA analyses in the main paper. At the same time, regression results do not show significant effects of both unpredictable ($p = 0.23$) and predictable ($p = 0.62$) *experienced* electrical stimulation on choice behavior. Specifically, we do not observe a significant effect of, or interaction with, both unpredictable and predictable electrical stimulation. Together, regression results support and extend the results reported in the main paper. Specifically, threat of shock remains a significant predictor of choice behavior ($p < 0.001$) in the trust and the non-social control game, even when controlling for the presence of actually experienced electrical stimulation. Furthermore, we show that the experience of shock does not significantly impact behavior in any of our analyses, indicating that the *expectancy* of shock, not the *experience* of shock, significantly impacted choice behavior.

Appendix

D Impulsivity, sensation seeking and negative emotionality relate to distinct morphometric brain features

Friederike Meyer, Jan B. Engelmann, Ernst Fehr, Christian C. Ruff

[1]Laboratory for Social and Neural Systems Research (SNS-Lab), Department of Economics, University of Zurich, Zurich, Switzerland

[2]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

D.1 Abstract

Individual differences in decision-making are often pragmatically assessed in healthy and psychiatric populations via personality traits that relate to real-life decision-making. Putative neurobiological bases of such personality traits have been identified in MRI studies reporting systematic relationships with stable (i.e., structural) as well as dynamic (i.e., functional) brain properties. However, the conceptual overlap of the psychological constructs used for personality assessment makes it difficult to draw firm conclusions regarding brain properties that uniquely predispose an individual towards a particular decision style. Here we employ voxel-based morphometry (VBM) to derive brain morphometric correlates of three personality dimensions relevant for decision-making: Impulsivity, sensation seeking and negative emotionality. These dimensions were identified by means of a factor analysis (FA), which was employed in order to address the possible colinearity of a broad range of personality characteristics associated with decision-making. With this approach, we yielded an improved assessment of the variance uniquely associated with the above-mentioned dimensions. Our VBM analysis showed that increased impulsivity and sensation seeking were uniquely associated with reduced grey matter (GM) volume in medial orbitofrontal (mOFC) and anterior cingulate cortex (ACC), respectively. Negative emotionality showed a unique negative relationship with GM volume in sections of the dorsolateral prefrontal (DLPFC) and inferior parietal cortex. Our results suggest that nonclinical individual differences in brain morphometry may constitute a neurobiological predisposition for impulsivity, sensation seeking, and negative emotionality, which may modify decision-making dependent on contextual and cognitive demands. Our findings also demonstrate that neurobiological information may be usefully employed in psychiatric research, since aberrant personality features and their modifying influence on behavior are hallmarks of psychiatric disorders.

D.2 Introduction

Individual differences in decision-making have been validated in different dimensions and their importance for explaining healthy as well as pathological behavior is commonly acknowledged (Der-Avakian & Markou, 2012; Lauriola, Panno, Levin, & Lejuez, 2013; Steinberg et al., 2008). Differences in observed choice behavior are systematically

associated with variation in self-reported personality dimensions such as negative emotionality, risk attitudes, impulsivity, and sensation seeking (Lauriola & Levin, 2001; Lauriola et al., 2013; Lauriola, Panno, Levin, & Lejuez, 2014). Many researchers in applied settings therefore pragmatically measure individual differences in decision-making with questionnaire measures of those personality dimensions, as it is non-trivial to use economic game paradigms (e.g., Jimura, Chushak, & Braver, 2013; Pine, Shiner, Seymour, & Dolan, 2010) for assessment of real-life decision-making styles (e.g., in routine clinical settings).

Previous studies suggest that individual personality differences may be rooted in biological factors such as differences in brain structure and function (Diekhof et al., 2012; Morishima, Schunk, Bruhin, Ruff, & Fehr, 2012). In the current study, we were especially interested in the neuroanatomical bases of the personality concepts such as impulsivity, sensation seeking, and negative emotionality, which show well-established correlations with everyday decision-making and are often used to quantify differences in decision-making (Ersche et al., 2012; Lauriola et al., 2013; Robbins, Gillan, Smith, de Wit, & Ersche, 2012; Wittmann & Paulus, 2008). Understanding the neurobiological sources of individual variation in these personality traits thus bears considerable value for research on clinical disorders, since aberrant degrees of impulsivity, sensation seeking, and negative emotionality are often reported for psychiatric disorders such as addiction, attention deficit hyperactivity disorder, as well as mood and anxiety disorders (e.g., Giorgetta et al., 2012; Lee, 2013; Paulus & Yu, 2012). Brain morphometric correlates of variation in these personality dimensions may therefore provide important information on possible neurocognitive endophenotypes of different subtypes of psychiatric disorders (Robbins et al., 2012).

Although studies have begun to identify neuroanatomical correlates of individual differences in impulsivity, sensation seeking, and negative emotionality (Matsuo et al., 2009; Schilling et al., 2012; Spampinato, Wood, De Simone, & Grafman, 2009), it is currently difficult to draw conclusions about the unique association between these personality constructs and brain morphometry. This is because most studies only investigated the association between a single personality construct and brain structure. Possible overlaps and colinearity between the different personality constructs (Whiteside & Lynam, 2001) are often not explicitly considered, so that many results may at least partially reflect the effects of confounding personality dimensions that were not concurrently assessed. Moreover, this issue is aggravated by the fact that it is hard to find a canonical conceptualization of the different personality dimensions. Competing theories use

overlapping concepts, and similar characteristics are defined with different names and slightly different meanings in multiple models. For instance, impulsivity is defined as a predisposition to choose unduly risky but enticing behavior paired with behavioral disinhibition (Daruma & Barnes, 1993), while sensation seeking describes a general tendency to seek novel, intensive emotional and highly stimulating experiences (Zuckerman, Eysenck, & J, 1978). It is evidently hard to adequately distinguish these concepts on purely theoretical grounds, even though some assessment instruments acknowledge the possible overlap and multidimensionality of these two constructs (Zuckerman, 1996b). Moreover, negative emotionality is defined as the tendency to experience unpleasant emotional states such as fear, nervous tension or anger (Hicks & Patrick, 2006). This concept is therefore related to emotional reactivity, which refers to the experience and magnitude of emotional arousal in terms of both negative and positive emotions (Glenn, Blumenthal, Klonsky, & Hajcak, 2011; Karrass et al., 2006).

In line with these conceptual difficulties of personality measurement, previous studies have not converged on distinct morphometric brain measures that uniquely relate to impulsivity, sensation seeking, and negative emotionality (e.g., Cho et al., 2013; Matsuo et al., 2009; Spampinato et al., 2009). The reasons for the inconsistent results yielded by the different studies (Cremers et al., 2011; DeYoung et al., 2010; Hu et al., 2011) are still debated, but are most likely attributable to insufficient sample sizes and differences in psychometric measures used to assess the same personality construct.

In the current study, we therefore employed a more comprehensive assessment approach that allows identification of unique variance in the relation between personality and brain morphometry. We acquired multiple questionnaire measures of personality characteristics traditionally found to be associated with individual differences in impulsivity, sensation seeking, and negative emotionality from 80 healthy student subjects, and submitted these to FA. The FA was performed in order to reduce redundancy between the multiple measures and produce a multidimensional assessment of personality patterns relevant for decision-making. These personality patterns were combined with VBM in order to investigate the neuroanatomical basis of the identified personality dimensions. This approach enabled us to identify associations between brain anatomy and specific personality traits while at the same time controlling for the influence of other choice-relevant personality characteristics.

D.3 Materials and methods

Subjects

T1-weighted images from 80 healthy subjects (63 male, mean age [SEM] = 22.23 [0.2]) were used in the VBM analysis. The structural brain scans as well as the questionnaire set were acquired in the context of two fMRI studies (Engelmann, Meyer, Fehr, & Ruff, 2015; Engelmann, Meyer, Ruff, & Fehr, submitted for publication). All procedures were approved by the ethics committee of the canton Zürich, Switzerland, and written informed consent was obtained from all participants.

Questionnaire measures

Questionnaires included in the current study were selected based on their hypothesized relevance for risky decision-making.

The multidimensional Neuroticism-Extraversion-Openness Five-Factor-Inventory (NEO-FFI; Borkenau & Ostendorf, 1993) was used to measure individual differences in the personality dimensions Neuroticism, Extraversion, Openness, Agreeableness and Conscientiousness, which have previously been shown to modulate risky choices (Capra, Jiang, Engelmann, & Berns, 2013). Multiple dimensions of impulsivity and sensation seeking were measured using the Barratt Impulsiveness scale (BIS-11; Patton, Stanford, & Barratt, 1995) and Zuckerman's Sensation Seeking Scale (Zuckerman, 1996a). Negative emotionality was assessed on multiple dimensions as well. Temporary and stable qualities of anxiety were assessed via the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). The Beck Depression Inventory (BDI; Hautzinger, Bailer, Worall, & Keller, 1995) was employed to measure depressive symptoms and severity. Different facets of risk-taking attitudes were assessed via the Domain-Specific-Risk-Taking Scale (DOSPERT; Weber, Blais, & Betz, 2002).

Factor analysis

To account for possible redundancy and colinearity of the 20 subscales that were employed, the item-level responses were scrutinized for underlying patterns via factor analytic procedures using Stata (StataCorp LP, College Station, Texas, USA). We applied Principal Factor analysis and a varimax rotation of the matrix of loadings in order to achieve orthogonal factors.

The factorability of the variable set was assessed with the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (MSA) and Bartlett's test of sphericity. A parallel analysis (Horn, 1965) was performed on the data set in order to identify the appropriate number of factors. Furthermore, the result of this parallel analysis was confirmed by use of a scree test (Cattell, 1966) of the eigenvalues plotted against the factor numbers. According to the Kaiser Criterion, we considered factors with eigenvalues < 1 as substantial (Kaiser, 1960). After identification of clearly defined and interpretable factors (item loadings > 0.40 were considered to be relevant factor loadings), we saved the subjects' loadings on these factors in the form of Bartlett factor scores. These factor scores were used to analyze the structural brain data for systematic relationships between personality measures and brain structure (see below).

MRI acquisition and VBM preprocessing

A Philips Achieva whole-body MR Scanner (Philips Medical Systems, Best, The Netherlands) equipped with an 8-channel Philips sensitivity encoded (SENSE) head coil was used for image acquisition. High-resolution structural T1-weighted 3D-TFE (3D-turbo fast echo) images (TR = 7500ms; TE = 3.5 ms; FA = 8; FOV 250 × 250 mm; voxel size 1.04 × 1.04 × 0.6 mm; 301 sagittal slices) were acquired for each participant.

We used SPM8 (http://www.fil.ion.ucl.ac.uk/spm) for MRI data preprocessing and analysis. Segmentation of native space images into gray and white matter and cerebral spinal fluid was conducted with the New Segment Toolbox (Ashburner & Friston, 2005). The segmented gray and white matter images (GM and WM, respectively) were aligned and warped to template space, followed by the resampling to 1.5 mm isotropic voxels. Inter-subject alignment was improved by registration of the resliced GM and WM images to a subject-specific template using the DARTEL Toolbox (Ashburner, 2007). After the normalization of both GM and WM images to MNI space (using the normalization function of the DARTEL Toolbox), the images were smoothed with a Gaussian kernel of full-width half-maximum of 8 mm.

ROI definition

Statistical analyses of the relationship between the personality dimensions identified by the FA (see "Results") and brain structure were performed for predefined ROIs. These anatomical regions were based on a-priori hypotheses derived from previous fMRI, VBM and MRS studies investigating functional and structural neural correlates of the personality

components identified in the current factor analysis (see below). Inference in the ROI analyses was performed using small volume correction (SVC; family-wise-error (FWE)-corrected at $p < 0.05$) for anatomical masks of these brain regions, which were created in the WFU Pickatlas (http://www.fmri.wfubmc.edu) using automated anatomical labeling (Tzourio-Mazoyer et al., 2002). For exploratory purposes, we also conducted whole-brain analyses corrected for multiple comparisons (FWE-corrected at $p < 0.05$). The following ROIs were selected based on published findings of associations with the relevant constructs:

Impulsivity/Risk-taking (IRT; factor 1): Anterior cingulate cortex (ACC) and ventral striatum (Cho et al., 2013; Cohen, Schoene-Bake, Elger, & Weber, 2009; Gallinat et al., 2007; Matsuo et al., 2009; Schilling et al., 2012).

Negative Emotionality (NE; factor 2): Amygdala, insula (Paulus & Stein, 2006), hippocampus (Hamilton et al., 2012; Hasler, Drevets, Manji, & Charney, 2004; Pitman, Shin, & Rauch, 2001; Yamasue et al., 2008), dorsolateral prefrontal cortex (DLPFC) (Bishop, 2009; Fitzgerald et al., 2006; Hamilton et al., 2012; Spampinato et al., 2009) and inferior parietal cortex (BA 40) (Schutter & van Honk, 2005; Schutter, Van Honk, Koppeschaar, & Kahn, 2002). The DLPFC mask was limited to BA9 intersected with middle frontal gyrus in order to restrict our analysis to a NE-relevant DLPFC subregion (Fitzgerald et al., 2006).

Sensation seeking (SS; factor 3): Medial orbitofrontal cortex (mOFC), inferior frontal cortex, anterior cingulate cortex, and striatum (Cohen et al., 2009; Gardini, Cloninger, & Venneri, 2009; Joseph, Liu, Jiang, Lynam, & Kelly, 2009; Jupp & Dalley, 2013; Van Schuerbeek, Baeken, De Raedt, De Mey, & Luypaert, 2011). The right and left mOFC masks combined medial orbitofrontal cortex and gyrus rectus masks from the WFU PickAtlas in order to cover the whole medial OFC region of interest (in line with Lim, O'Doherty, & Rangel, 2011).

Statistical parametric mapping of segmented tissue density

In order to investigate the covariation between grey matter (GM) tissue density and the factorized personality parameters, the preprocessed segmented images were entered into a multiple regression model in SPM8. The calculated factor scores were entered as covariates of interest for each subject. Age and gender of each subject were included into the GLM as nuisance covariates to account for variance due to these variables. Each subject's GM image was proportionally scaled by his individual total intracranial volume (TIV) using SPM's proportional Global Scaling option (as also used in, e.g., Morishima et al., 2012). A

GM mask was calculated by thresholding the mean of all subjects' smoothed DARTEL-normalized GM images at a signal intensity of 0.2 in order to restrict the search volume to GM. The resulting image was used as explicit mask in the multiple regression analysis.

D.4 Results

Summary statistics

A summary of the calculated questionnaire scores can be found in Table D-I. The individual scores yielded by each subject were used as items in the factor analysis.

Table D-I: Summary statistics (mean and standard deviation [SD]) of personality questionnaire set

Variable	Mean	SD	Variable	Mean	SD
STAI X1	36.41	6.3	STAI X2	38.85	8.76
NEO NE	19.1	7.09	DOSPERT ET	18.94	5.36
NEO EX	30.14	6.25	DOSPERT FI	17.15	5.02
NEO OP	34.1	6.74	DOSPERT HE	22.64	5.4
NEO AG	31.36	5.66	DOSPERT RE	26.06	6.87
NEO CO	28.81	8.22	DOSPERT SO	28.96	3.54
SSS TA	7.73	2.44	BIS-11 AT	16.35	3.3
SSS DI	6.36	2.36	BIS-11 MT	23.65	4.72
SSS EX	6.88	1.94	BIS-11 NP	24.3	4.38
SSS BD	4.76	1.86	BDI	6.39	4.59

Note: BDI: Beck Depression Inventory; STAI: X1 = state anxiety; X2 = trait anxiety, BIS-11 subscales: AT = attention, MT = motor; NP = non-planning; DOSPERT subscales: ET = ethical, FI = financial, HE = health, RE = recreational, SO = social, NEO (NEO-FFI): NE = neuroticism; E = extraversion, O = openness; AG = agreeableness, CO = conscientiousness, SSS subscales: TA = thrill and adventure seeking, DI = disinhibition, Ex = experience seeking; BD = boredom susceptibility

Factor analysis

Estimation of the KMO criterion for the single variables contained in the FA (performed on all subscales) revealed low MSA for the subscale "financial risk-taking" of the DOSPERT (KMO = 0.35) and the subscale "agreeableness" of the NEO-FFI (KMO = 0.35). These two subscales were thus excluded from the FA.

A Principal Factor analysis with a varimax (i.e., orthogonal) rotation of the remaining 18 questionnaire subscales was conducted on the data gathered from 80 participants. The validity of this approach was confirmed by the KMO criterion (KMO = .72 for the current data set) and the finding that Bartlett's test of sphericity rejected the null hypothesis of no intercorrelation between the variables (χ^2 (153) = 777.57, $p < 0.0001$).

Both parallel analysis and scree test suggested a three-factor solution for the revised data set, which was partly consistent with the Kaiser criterion suggesting four factors. Taking into account that the third factor built the "elbow" of the eigenvalue function and that the fourth factor (eigenvalue = 1.03) did not explain a substantial proportion of variance, the three-factor solution was chosen (Figure D-1). Approximately 82% of the variance found in the variable set was explained by this three-factor model (Table D-II). Most importantly, this solution yielded interpretable and theory-consistent factors.

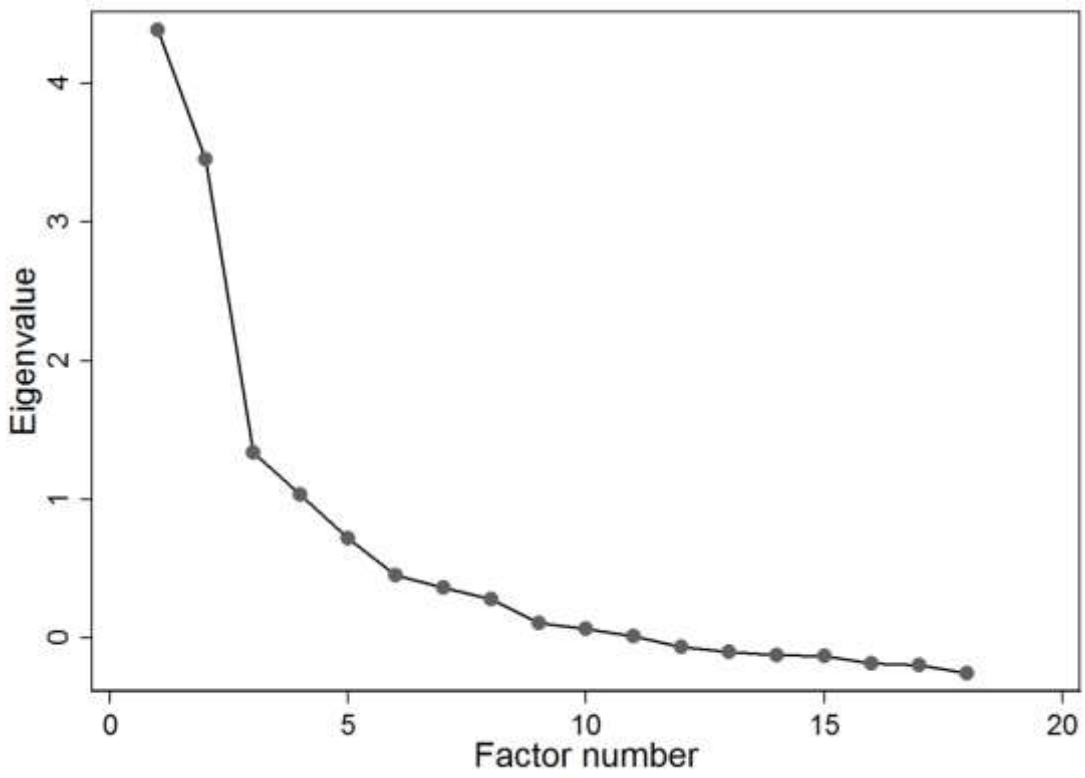

Figure D-1: Scree plot of eigenvalues plotted against number of factors. Scree plot combined with Kaiser Criterion (eigenvalues: factor 1 = 4.41, factor 2 = 3.48, factor 3 = 1.36, factor 4 = 1.03) suggested a 3-factor solution for the questionnaire set.

Table D-II: Variance of data set explained by three-factor model

Factor	Variance	Proportion	Cumulative
Factor 1	3.37	0.3	0.3
Factor 2	3.32	0.3	0.6
Factor 3	2.48	0.22	0.82

Note: Table contains variance accounted for by each factor, proportion of variance accounted for by each factor, and cumulative proportion of variance accounted for by the factor plus previous factors.

Marker variables of the first factor were *impulsivity and risk-taking related* items (short: IRT), as we observed strong positive loadings of the BIS-11 subscale non-planning, the DOSPERT health risk-taking subscale and a negative loading of the NEO conscientiousness subscale onto this factor. The second factor reflected multiple dimensions of *negative emotionality* (NE), as marker variables of this factor were the BDI, quality of long-term anxiety (STAI-T) and the NEO subscale neuroticism. The third factor explained facets of sensation seeking (SS), as the DOSPERT subscale "recreational risk-taking" and the SSS "thrill and adventure seeking" subscales loaded onto this factor (Table D-III).

Table D-III: Factor loadings for the three-factor model

Variable	Factor 1	Factor 2	Factor 3	Uniqueness
BDI		0.81		0.33
STAIT X1		0.58		0.66
STAIT X2		0.89		0.19
BIS-11 AT	0.44	0.46		0.59
BIS-11 MT	0.62			0.47
BIS-11 NP	0.79			0.37
DOSPERT ET	0.67			0.53
DOSPERT HE	0.73			0.39
DOSPERT RE			0.71	0.44
DOSPERT SO				0.81
NEO NR		0.87		0.22
NEO EX			0.56	0.52
NEO OP			0.57	0.59
NEO CO	-0.71			0.4
SSS TA			0.69	0.48
SSS DI	0.42			0.7
SSS EX	0.53		0.55	0.41
SSS BD			0.46	0.72

Note: Factor 1 was identified as reflecting impulsivity- and risk-taking, factor 2 comprises different aspects of negative emotionality (NE), whereas factor 3 reflects facets of sensation seeking (SS).

VBM: Brain-personality associations

We performed a multiple regression analysis of the structural brain data on the complete set of factors identified by the factor analysis (plus nuisance covariates age and gender). This approach was taken in order to reveal the unique variance accounted for by a personality factor of interest, while at the same time controlling for the potentially confounding effects of the other personality factors.

Morphometric brain correlates of impulsivity and risk-taking (IRT; factor 1)

We found that IRT showed a systematic negative correlation with regional GM volume in the anterior mOFC. That is, higher degrees of self-reported facets of impulsivity were associated with decreased GM volume in both left mOFC ($x = -6$, $y = 51$, $z = -24$; $T = 3.54$, $p < 0.05$, SVC) and right mOFC ($x = 9$, $y = 45$, $z = -21$, $T = 3.71$, $p < 0.05$, SVC). This relation was highly specific. First, IRT was not significantly correlated with GM volume in

the other ROIs (amygdala, insula, striatum, DLPFC, inferior parietal cortex, inferior frontal cortex, ACC). Second, this relation was specific to IRT as no significant relation between mOFC GM volume and SS or NE was found (i.e., no significant voxels in these ROIs at $p < 0.05$, SVC; Figure D-2, Table D-IV). Notably, these findings complement previous reports on morphometric brain correlates of impulsivity and risk-taking. Although both Cho et al. (2013) and Matsuo et al. (2009) found that OFC GM volume is related to impulsive tendencies, the specifics of the observed relationship between impulsivity and brain structure differed between these researchers. That is, whereas the former authors found that higher BIS-11-measured non-planning, attention-related and total score impulsivity were associated with higher *m*OFC GM volume, the latter reported an inverse correlation of BIS-11-measured non-planning, total score and motor impulsivity and *lateral* OFC GM volume. We might speculate that study characteristics such as number of subjects, different measures of personality assessment or inclusion of nuisance variables led to these divergent results, as VBM is a very susceptible measure (Hu et al., 2011).

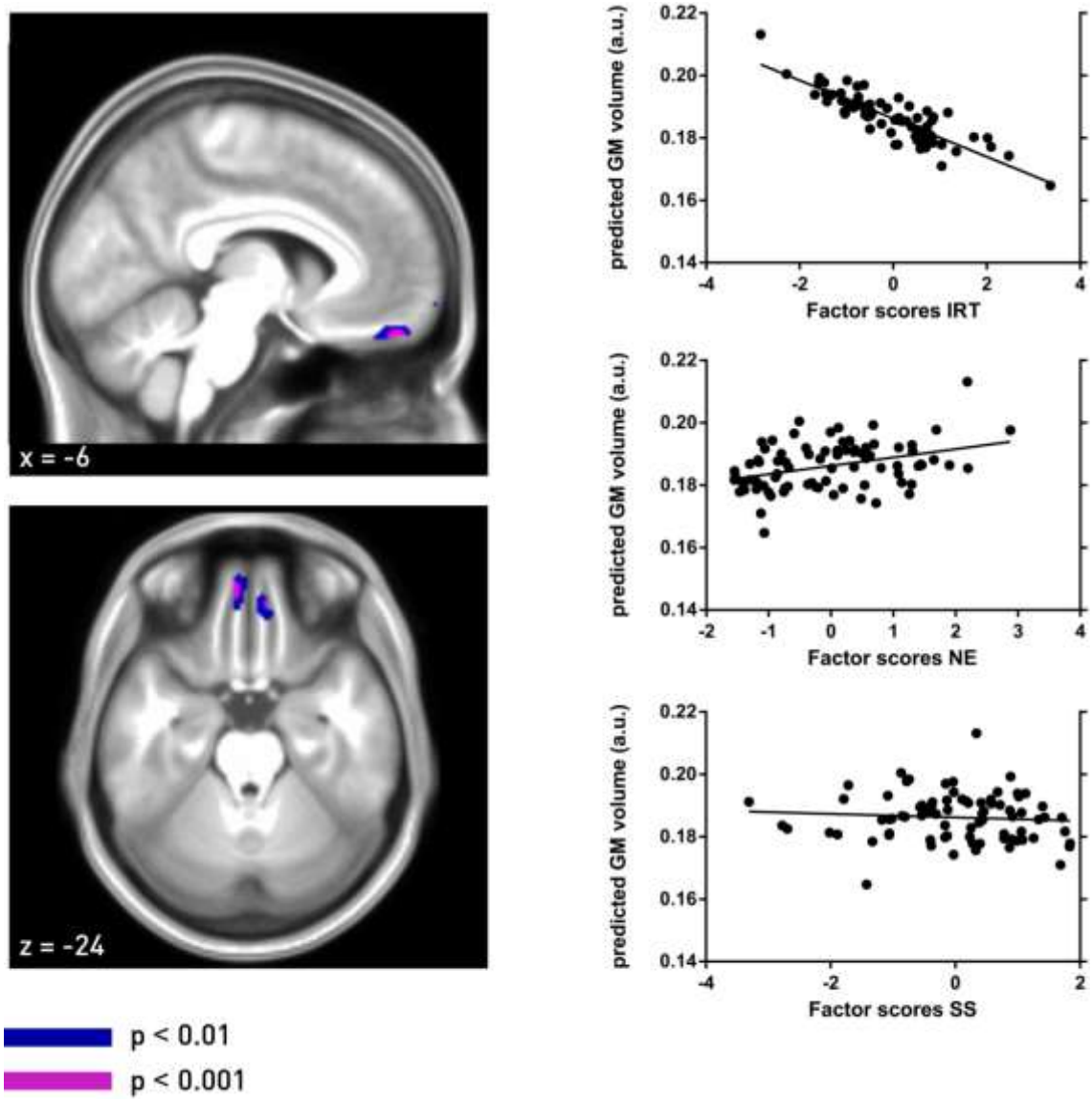

Figure D-2: Impulsivity and Risk-Taking (IRT) is associated with decreased GM volume in mOFC. The brain slices (right panel) show the SPM of the negative regression contrast (thresholded at $p < 0.001$ and $p > 0.01$ for display purposes) of the scaled GM volume (i.e., normalized by TIV) on individual IRT scores. For illustration purposes, the scatter plots (right panel) show the predicted GM volume estimates for each participant extracted from left mOFC peak voxel (x = -6, y = 51, z = -24; T = 3.54, $p < 0.05$, SVC) plotted against all three factor scores. Regression lines are shown to visualize the direction of the effects. A systematic relationship is evident for IRT (top plot) but not for NE and SS (middle and bottom plot, respectively).

Table D-IV: Negative relationship between IRT and left mOFC GM volume identified by multiple regression analysis

Variables	Coefficient	SE	T	p
IRT	-0.00638***	0.0018	-3.542	0.001
NE	0.00188	0.00195	0.962	0.339
SS	-0.000869	0.00176	-0.493	0.623
Age	-0.000833	0.00098	-0.850	0.398
Gender	0.0051	0.00498	1.021	0.311
Constant	0.204***	0.022	9.262	0.000
R^2	0.177			

Note: We controlled for individual differences in TIV by dividing GM volume at the selected voxel (x = -6, y = 51, z = -24) by TIV. In addition to the personality parameters of interest, we also include age and gender as regressors of no interest in our model. Only impulsivity shows a significantly negative association with mOFC GM volume. SE: standard error. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Morphometric brain correlates of negative emotionality (NE, factor 2)

We identified a negative relationship between NE and regional left mid-DLPFC, specifically left BA9. Subjects with higher scores on the NE factor had lower GM volume in this part of the prefrontal cortex (x_1 = -42, y_1 = 23, z_1= 34, T = 3.38; $p < 0.05$, SVC; x_2 = -41, y_2 = 33, z_2 = 39, T = 3.36, $p < 0.05$, SVC, Figure D-3a, Table D-V). Note that this region corresponds well with the portion of BA 9 that was shown to correlate with anxiety and depressive symptoms by Fitzgerald et al. (2006). Moreover, NE correlated negatively with GM volume in the right inferior parietal cortex (IPC; x = 46, y = -48, z = 58, T = 4.08, $p < 0.05$, SVC, Figure D-3b, Table D-VI). Again, the relation between NE and GM volume in left DLPFC and IPC was highly specific: We did not find any other significant relationships between NE and GM volume in the other ROIs (amygdala, insula, striatum, mOFC, inferior frontal cortex, ACC). Moreover, GM volume in both regions was clearly not correlated with individual IRT or SS at the applied statistical threshold of $p < 0.05$, SVC, confirming the selectivity of the association between NE and morphometry in the DLPFC and inferior parietal cortex

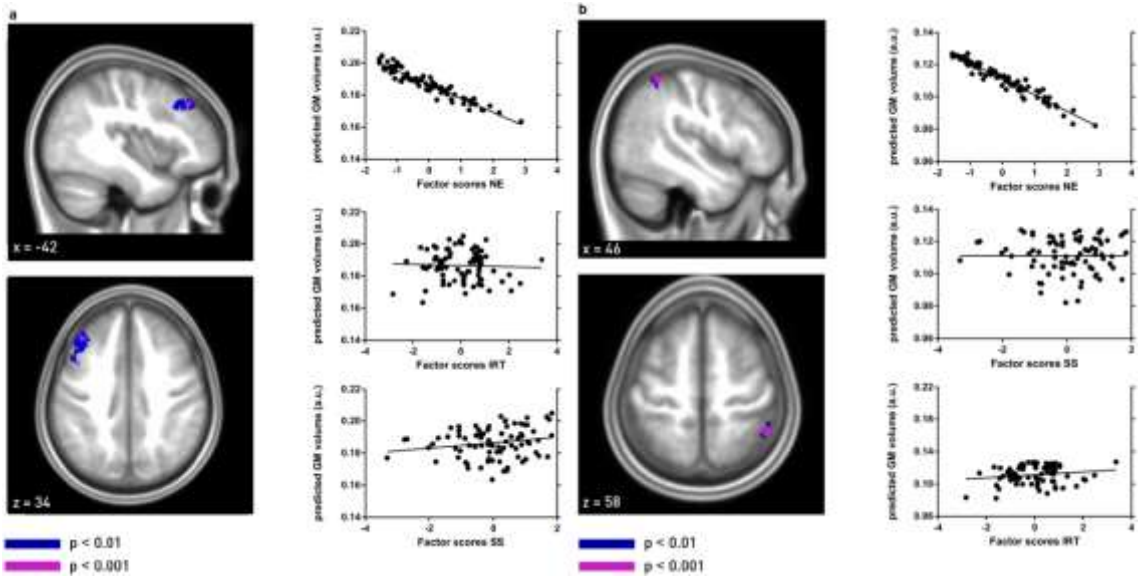

Figure D-3: (a) Negative emotionality (NE) correlates with reduced GM volume in the left DLPFC. The brain images (left side) show thresholded SPMs (at p<.001 and p<.01 for display purposes) of the negative regression weight for scaled GM volume (i.e., normalized by TIV) on individual NE scores. The scatter plots on the right show the predicted GM volume estimates in the DLPFC peak voxel (x = -42, y = 23, z = 34; T = 3.38, p < 0.05, SVC) plotted against all three factor scores. GM volume in left DLPFC is negatively related to higher NE scores (top plot), but not to IRT or SS (middle and bottom plot, respectively). (b): NE correlates with reduced GM volume in the right inferior parietal cortex (IPC). The brain images on the left show thresholded SPMs (at p < 0.001 and p < 0.01 for display purposes) of the negative regression weight for scaled GM volume (i.e., normalized by TIV) on individual NE scores. For illustration purposes, the scatter plots on the right show the predicted GM volume estimates in the right IPC peak voxel (x = 46, y = -48, z = 58; T = 4.08, p < 0.05, SVC) plotted against all three factor scores. GM volume in right IPC is negatively related to higher NE scores (top plot), but not to IRT or SS (middle and bottom plot, respectively). Regression lines are displayed to visualize the direction the effects.

Table D-V: Negative correlation between NE and local DLPFC GM volume identified by multiple regression analysis

Variables	Coefficient	SE	T	p
IRT	-0.0004	0.00246	-0.162	0.871
NE	-0.00904***	0.00267	-3.380	0.001
SS	0.00185	0.00241	0.766	0.446
Age	0.00075	0.00134	0.562	0.576
Gender	0.00166	0.00681	0.244	0.808
Constant	0.169***	0.0301	5.631	0.000001
R^2	0.151			

Note: We controlled for individual differences in TIV by dividing GM volume at the selected voxel (x = -42, y = 23, z = 34) by TIV. In addition to the personality parameters of interest, we also include age and gender as regressors of no interest in our model. We find a significantly negative relationship between local DLPFC GM volume and NE, whereas no other personality dimension shows a significant association with DLPFC GM volume. SE: standard error. *** p < 0.01, ** p < 0.05, * p < 0.1

Table D-VI: Negative correlation between NE and local inferior parietal cortex GM identified by multiple regression analysis

Variables	Coefficient	SE	T	p
IRT	0.0015	0.00208	0.730	0.468
NE	-0.0092***	0.00225	-4.085	0.0001
SS	0.0005	0.00203	0.247	0.805
Age	-0.0003	0.00113	-0.267	0.790
Gender	-0.0051	0.00574	-0.896	0.373
Constant	0.119***	0.0253	4.692	0.00001
R^2	0.245			

Note: We controlled for individual differences in TIV by dividing GM volume at the selected voxel (x = 46, y = -48, z = 58) by TIV. In addition to the personality parameters of interest, we also include age and gender as regressors of no interest in our model. Increased NE scores were correlated with decreased GM volume in this brain region. No other personality dimension investigated showed significant correlations with inferior parietal cortex GM volume. SE: standard error *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Morphometric brain correlates of sensation seeking (SS; factor 3)

For the third factor, SS, we found a significant negative correlation with regional GM volume in bilateral ACC ($x_R = 11$, $y_R = 41$, $z_R = 28$; $T = 4.33$, $p < 0.05$, SVC; $x_L = -2$, $y_L = 8$, $z_L = 30$, $T = 3.5$, $p < 0.05$, SVC). This finding corresponds with the hypothesized role of the ACC in motivation and drive and is in accordance with previous studies on ACC function for motivated behavior (e.g., Gallinat et al., 2007). We did not find significant correlations with SS for GM volume of the other ROIs (amygdala, insula, striatum, mOFC, DLPFC, inferior parietal cortex, inferior frontal gyrus) at $p < 0.05$, SVC. Moreover, neither IRT nor NE showed significant correlation with regional ACC GM volume (Figure D-4, Table D-VII).

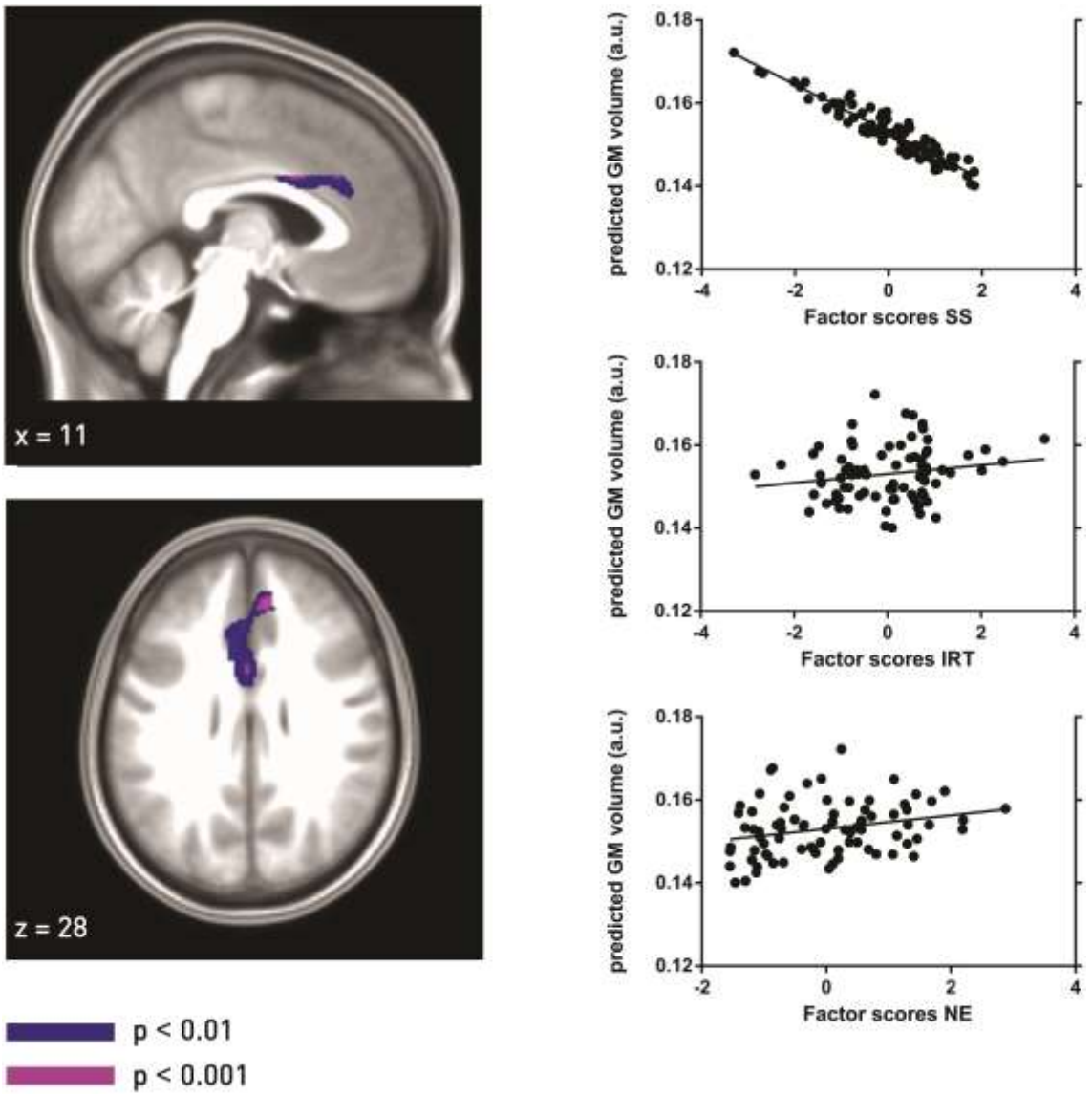

Figure D-4: Sensation Seeking (SS) is associated with reduced GM volume in the anterior cingulate cortex (ACC). The brain images on the left show thresholded SPMs (at $p < 0.001$ and $p < 0.01$ for display purposes) of the negative regression weight for scaled GM volume (i.e., normalized by TIV) on individual SS scores. The observed significant relationship between SS and GM in ACC was highly specific as neither IRT nor SE GM was systematically associated with ACC GM volume. For illustration purposes, the scatter plots on the right show the predicted GM volume estimates in the right ACC peak voxel (x = 11, y = 41, z = 28¸ T = 4.33, $p < 0.05$, SVC) plotted against all three factor scores. Regression lines are also displayed to visualize the direction the effects. GM volume in ACC is negatively related to Higher SS scores (top plot), but not to IRT or NE (middle and bottom plot, respectively).

Table D-VII: Multiple regression analysis on the relationship between local ACC volume and personality parameters.

Variables	Coefficient	SE	T	p
IRT	0.000744	0.00135	0.550	0.584
NE	0.0019	0.00147	1.291	0.201
SS	-0.00573***	0.00132	-4.328	0.000046
Age	0.00009	0.00736	0.125	0.901
Gender	-0.0001	0.00374	-0.042	0.966
Constant	0.151***	0.0165	9.146	0.000
R^2	0.221			

Note: We controlled for individual differences in TIV by dividing GM volume at the selected voxel (x = 11, y = 41, z = 28) by TIV. In addition to the personality parameters of interest, we also include age and gender as regressors of no interest in our model. We find a significantly negative relationship between local ACC GM volume and SS, whereas no other personality dimension shows a significant association with ACC GM volume. SE: standard error. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

D.5 Discussion

Our study captured the overlap and redundancy of multiple affective and personality dimensions relevant for daily decision-making by means of a factor analysis in order to identify unique anatomical brain correlates of these traits. We demonstrated systematic and specific associations between these personality dimensions and brain structure: A systematic negative correlation was found between impulsivity and regional mOFC GM volume, whereas higher degrees of negative emotionality were associated with lower regional GM volume in the DLPFC and inferior parietal cortex. Finally, sensation seeking was negatively correlated with regional GM volume in the ACC. Importantly, these brain regions have been routinely found to relate to various aspects of decision-making in neuroeconomic research (Rushworth, Noonan, Boorman, Walton, & Behrens, 2011; Walton, Behrens, Noonan, & Rushworth, 2011). Our results thereby suggest that brain structure in these areas may predispose individuals towards particular affective and personality traits as well as a specific decision-making style. Furthermore, this neuroanatomical overlap might offer an explanation for why distorted decision-making processes in psychiatric disorders (e.g., depression, anxiety disorders, ADHD, addiction, bipolar disorder, schizophrenia) occur jointly with comorbid aberrations in the personality domains identified in the present study (e.g., Ersche et al., 2012; Reddy et al., 2014).

Our present VBM investigation demonstrates the need to distinguish between impulsivity and sensation seeking, since these constructs were identified as distinct on both the behavioral and the structural brain level, with unique relations between impulsivity and regional mOFC GM volume versus a unique association between SS and regional ACC GM volume. This dissociation underlines that the two concepts imply fundamentally distinct aspects – despite their colinearity and overlap – which can be identified if the two concepts are acquired simultaneously and included in the same model, as done in the present study. This procedure may improve the specificity of findings on brain structure-behavior relations. For instance, previous studies (Matsuo et al., 2009) have reported a significant correlation between impulsivity and ACC GM volume (as well as lateral OFC GM volume). Our new results suggest that these previous findings may reflect the variance that is shared between impulsivity and sensation seeking, which has not been controlled for in previous investigations such as the above mentioned, and, instead, show an association between mOFC GM volume and impulsivity.

The observed relationship between regional ACC GM volume and SS is consistent with functional MRI results by Joseph et al. (2009) who report less ACC responsiveness to highly arousing stimuli in high relative to low sensation seekers. The authors interpreted their findings as a more active approach system (relative to inhibitory system) in high versus low sensation seekers. In addition, Gallinat et al. (2007) found that increased SSS-V scores correlated negatively with glutamate level in ACC. Our new results might bridge the gap between these functional and structural brain correlates of sensation seeking, by demonstrating a congruency between neural activity and brain morphometry associated with this personality trait. Regional ACC GM volume may be one structural index of a specific neurobiological basis for sensation seeking in this brain structure, which interacts with extrinsic and intrinsic factors in order to determine brain activity and thereby individual behavior. Our findings indicate a specific link of ACC GM volume and sensation seeking, as we did not observe a significant correlation between impulsivity and ACC volume (Matsuo et al., 2009).

With respect to impulsivity and risk-taking, our results point to the mOFC as a putative neuroanatomical basis of this trait, as we found a negative correlation between IRT and mOFC volume. This finding dovetails with observations of decreased resting-state activity in medial OFC in persons with Borderline personality disorder (BPD) compared to healthy controls, a disorder known for elevated impulsivity (Wolf et al., 2012). However, contrary to the expectations of Wolf et al. (2012) to find a negative relationship between

impulsivity and mOFC baseline perfusion, the authors observed a positive correlation between BIS-11 total impulsivity scores and mOFC blood flow. That is, higher impulsivity was associated with increased resting-state blood flow in this brain region. In order to explain this puzzling result, Wolf et al. (2012) hypothesized that this positive association between impulsivity and mOFC blood flow might be indicative of different levels of arousal or decreased intrinsic alertness in BPD relative to healthy control subjects. Furthermore, the authors suggested that the observed relationship might reflect different functional activation patterns during resting state versus (experimental) external stimulation. It should be noted, however, that the researchers calculated the correlation between impulsivity and mOFC blood flow across BPD subjects and healthy controls. This approach did not consider systematically different relationships between impulsivity and blood flow in mOFC in BPD versus healthy controls, which might represent a bias of the performed correlation analysis. Furthermore, studies on ADHD, a clinical condition known for aberrant impulsivity levels, found decreased reward value coding in mOFC correlated with behavioral impulsivity (Wilbertz et al., 2012), that is, deficient mOFC coding of motivational changes in reward value correlated negatively with experimentally measured impulsivity. This finding underscores the functional role of mOFC in impulsive behavior and complements our results on brain structure-behavior correlation.

However, inconsistent with our results, Cho et al. (2013) found a positive relationship between mOFC GM volume and BIS-11 measured impulsivity. In particular, the authors reported positive correlations of mOFC GM volume with BIS total score, non-planning as well as attention-related impulsivity. We might speculate that experimental factors such as the smaller number of subjects and the restriction of impulsivity measurement to BIS-11 in their study are reasons for the divergent results. Matsuo et al. (2009) reported an inverse relationship between *middle* OFC and BIS-11 measured impulsivity. Specifically, BIS-11 motor impulsivity negatively correlated with left OFC GM volume, while higher BIS-11 non-planning impulsivity scores were associated with smaller right OFC GM volume. These results clearly differ from the negative relationship between impulsivity and *medial* OFC morphometry found in the current study.

In a broader sense, the above outlined and to some degree inconsistent results point to the crucial and yet underinvestigated distinction between neurofunctional versus neuroanatomical patterns and their association with observed behavior. Future studies should thus address the relationship between impulsivity and OFC morphometry in a more

focused way than done until date in order to answer the open questions concerning the neurobiological basis of impulsivity raised by the conducted studies.

Regarding IRT and SS, our results add to previous neuroimaging and VBM studies that have pointed to the role of the mOFC in impulsive and risk-taking behavior, and they underline the specificity of this relationship and the essential distinction from sensation seeking on the behavioral and brain level.

Furthermore, negative emotionality was negatively associated with GM volume in the DLPFC and inferior parietal cortex. These results confer with previous studies that have revealed pathological activation of both DLPFC and parietal cortex in anxiety and depression (Pallanti & Bernardi, 2009; Pascual-Leone, Rubio, Pallardó, & Catalá, 1996; Vasic, Walter, Sambataro, & Wolf, 2009; Wolkenstein & Plewnia, 2013).

Furthermore, the finding of a negative association between regional DLPFC and inferior parietal GM volume with negative emotionality is consistent with the proposed *functional left prefrontal – right parietal* cortico-cortical depression circuit (Rosenberg et al., 2002; Schutter et al., 2002), which refers to a hypoactivation of left prefrontal and right parietal cortex in depression observed by these researchers. Our new results now extend these clinical findings, by showing that variations in morphometry of these brain structures also relate to pre-clinical individual variation in negative emotionality. This finding suggests that GM volume of these brain areas may possibly be used to identify endophenotypes of psychiatric disorders and in the context of dimensional diagnostics, but future studies should first further consolidate the association between brain structure and negative emotionality with a wider array of psychometric and physiological measurement instruments.

At a more general level, the results from the present study support the notion that personality factors can be important mediators of the heterogeneity commonly observed in economic and social choice (Capra et al., 2013). We show that personality factors relate specifically to decreasing GM volume in mOFC, DLPFC and ACC. These regions are routinely found to relate to various aspects of decision-making in neuroeconomics research, such as value computations and action selection (Rushworth et al., 2011; Walton et al., 2011). This neuroanatomical overlap between brain correlates of personality constructs and decision processes offers a possible explanation for why distorted decision-making processes observed in mood disorders (Der-Avakian & Markou, 2012; Huys, Pizzagalli, Bogdan, & Dayan, 2013; Paulus & Yu, 2012) commonly occur in conjunction with

comorbid aberrations in the personality domains identified in the present study (e.g., Ersche et al., 2012).

On a more critical note, the suggested possible links between brain function and structure described above clearly have to be taken with caution. Strong empirical evidence for precise structure-function relationships is largely missing until date (though see Morishima et al., 2012) and the theoretical basis for such links still need to be fully established (Koelsch, Skouras, & Jentschke, 2013; Wagner et al., 2008). Nevertheless, our current results suggest a congruency between functional responses and morphometric properties in the brain, as we found a negative relationship between brain morphometry and personality traits that are often observed to modify decision-related functional brain responses (Bishop, 2009; Joseph et al., 2009; Sripada, Gonzalez, Phan, & Liberzon, 2011). We thus speculate that differences in GM volume may account for the interindividual heterogeneity in baseline personality properties, which influence behavior depending on contextual and cognitive demands of the choice situation. However, this interpretation is clearly limited due to the correlative nature of the VBM approach, which prevents conclusions about causal mechanisms leading to the observed personality-morphometry associations. Is individual heterogeneity in personality traits caused by differences in brain morphometry or do differences in personality modify brain morphometry? While proper inferences about this may only be drawn from future longitudinal and interventional studies, previous demonstrations of genetic influences on the heritability of personality traits (Cloninger, Adolfsson, & Svrakic, 1996; Comings et al., 2000; Munafò & Flint, 2011) are consistent with the assumption that human personality may be at least partially determined by brain morphometry.

To conclude, our study elucidates the neurobiological basis of personality traits that modify individual decision-making. By simultaneously assessing several of these personality traits and accounting for their shared variance in the analysis, we were able to distinguish two commonly confounded concepts, impulsivity and sensation seeking, on the behavioral and neural level. This distinction might be especially important in the domain of diagnostics and treatment of addiction. It has been shown that impulsivity might be a risk factor for the development of a stimulant dependency, whereas increased sensation seeking seemed to be a consequence of the stimulant addiction (Ersche, Turton, Pradhan, Bullmore, & Robbins, 2010). Our investigation thus informs psychiatric research and may lead to new impulses for the diagnostics of clinical disorders. For example, it may support identifying endophenotypes of psychiatric disorders and supporting dimensional diagnostics). At a

more general level, our investigation suggests that research on the relation between brain structure and personality traits clearly benefits from an approach that explicitly measures the colinearity and conceptual overlap of competing personality constructs.

D.6 References

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, *38*(1), 95–113. doi:10.1016/j.neuroimage.2007.07.007

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–851. doi:10.1016/j.neuroimage.2005.02.018

Bishop, S. J. (2009). Trait anxiety and impoverished prefrontal control of attention. *Nature Neuroscience*, *12*(1), 92–98. doi:10.1038/nn.2242

Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa und McCrae*. Göttingen: Hogrefe.

Capra, C. M., Jiang, B., Engelmann, J. B., & Berns, G. S. (2013). Can Personality Type Explain Heterogeneity in Probability Distortions? *Journal of Neuroscience, Psychology, and Economics*, *6*(3), 151–166.

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, *1*(2), 245–276. doi:10.1207/s15327906mbr0102_10

Cho, S. S., Pellecchia, G., Aminian, K., Ray, N., Segura, B., Obeso, I., & Strafella, A. P. (2013). Morphometric Correlation of Impulsivity in Medial Prefrontal Cortex. *Brain Topography*, *26*(3), 479–487. doi:10.1007/s10548-012-0270-x

Cloninger, C. R., Adolfsson, R., & Svrakic, N. M. (1996). Mapping genes for human personality. *Nature Genetics*, *12*(1), 3–4. doi:10.1038/ng0196-3

Cohen, M. X., Schoene-Bake, J.-C., Elger, C. E., & Weber, B. (2009). Connectivity-based segregation of the human striatum predicts personality characteristics. *Nature Neuroscience*, *12*(1), 32–34. doi:10.1038/nn.2228

Comings, D., Gade-Andavolu, R., Gonzalez, N., Wu, S., Muhleman, D., Blake, H., … MacMurray, J. (2000). A multivariate analysis of 59 candidate genes in personality

traits: the temperament and character inventory. *Clinical Genetics*, *58*(5), 375–385. doi:10.1034/j.1399-0004.2000.580508.x

Cremers, H., van Tol, M.-J., Roelofs, K., Aleman, A., Zitman, F. G., van Buchem, M. A., … van der Wee, N. J. A. (2011). Extraversion is linked to volume of the orbitofrontal cortex and amygdala. *PloS One*, *6*(12), e28421. doi:10.1371/journal.pone.0028421

Daruma, J., & Barnes, P. (1993). A neurodevelopmental view of impulsivity and its relationship to the superfactors of personality. In W. McCown, J. Johnson, & M. Shure (Eds.), *The Impulsive Client: Theory, Research and Treatment* (Washington, DC.). American Psychological Association.

Der-Avakian, A., & Markou, A. (2012). The neurobiology of anhedonia and other reward-related deficits. *Trends in Neurosciences*, *35*(1), 68–77. doi:10.1016/j.tins.2011.11.005

DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., & Gray, J. R. (2010). Testing predictions from personality neuroscience. Brain structure and the big five. *Psychological Science*, *21*(6), 820–828. doi:10.1177/0956797610370159

Diekhof, E. K., Nerenberg, L., Falkai, P., Dechent, P., Baudewig, J., & Gruber, O. (2012). Impulsive personality and the ability to resist immediate reward: An fMRI study examining interindividual differences in the neural mechanisms underlying self-control. *Human Brain Mapping*, *33*(12), 2768–2784. doi:10.1002/hbm.21398

Engelmann, J.B., Meyer, F., Fehr, E., Ruff, C.C. (2015). Anticipatory anxiety disrupts neural valuation during risky choice. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(7), 3085–3099. doi:10.1523/JNEUROSCI.2880-14.2015

Engelmann, J. B., Meyer, F., Ruff, C. C., & Fehr, E. (submitted for publication). The neural circuitry of emotion-induced distortions of trust.

Ersche, K. D., Turton, A. J., Chamberlain, S. R., Müller, U., Bullmore, E. T., & Robbins, T. W. (2012). Cognitive dysfunction and anxious-impulsive personality traits are endophenotypes for drug dependence. *The American Journal of Psychiatry*, *169*(9), 926–936. doi:10.1176/appi.ajp.2012.11091421

Ersche, K. D., Turton, A. J., Pradhan, S., Bullmore, E. T., & Robbins, T. W. (2010). Drug addiction endophenotypes: impulsive versus sensation-seeking personality traits. *Biological Psychiatry*, *68*(8), 770–773. doi:10.1016/j.biopsych.2010.06.015

Fitzgerald, P. B., Oxley, T. J., Laird, A. R., Kulkarni, J., Egan, G. F., & Daskalakis, Z. J. (2006). An analysis of functional neuroimaging studies of dorsolateral prefrontal cortical activity in depression. *Psychiatry Research*, *148*(1), 33–45. doi:10.1016/j.pscychresns.2006.04.006

Gallinat, J., Kunz, D., Lang, U. E., Neu, P., Kassim, N., Kienast, T., … Bajbouj, M. (2007). Association between cerebral glutamate and human behaviour: the sensation seeking personality trait. *NeuroImage*, *34*(2), 671–678. doi:10.1016/j.neuroimage.2006.10.004

Gardini, S., Cloninger, C. R., & Venneri, A. (2009). Individual differences in personality traits reflect structural variance in specific brain regions. *Brain Research Bulletin*, *79*(5), 265–270. doi:10.1016/j.brainresbull.2009.03.005

Giorgetta, C., Grecucci, A., Zuanon, S., Perini, L., Balestrieri, M., Bonini, N., … Brambilla, P. (2012). Reduced risk-taking behavior as a trait feature of anxiety. *Emotion (Washington, D.C.)*, *12*(6), 1373–1383. doi:10.1037/a0029119

Glenn, C. R., Blumenthal, T. D., Klonsky, E. D., & Hajcak, G. (2011). Emotional reactivity in nonsuicidal self-injury: divergence between self-report and startle measures. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *80*(2), 166–170. doi:10.1016/j.ijpsycho.2011.02.016

Hamilton, J. P., Etkin, A., Furman, D. J., Lemus, M. G., Johnson, R. F., & Gotlib, I. H. (2012). Functional neuroimaging of major depressive disorder: a meta-analysis and new integration of base line activation and neural response data. *The American Journal of Psychiatry*, *169*(7), 693–703. doi:10.1176/appi.ajp.2012.11071105

Hasler, G., Drevets, W. C., Manji, H. K., & Charney, D. S. (2004). Discovering endophenotypes for major depression. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *29*(10), 1765–1781. doi:10.1038/sj.npp.1300506

Hautzinger, M., Bailer, M., Worall, H., & Keller, F. (1995). *Beck-Depressions-Inventar (BDI). Testhandbuch* (2nd ed.). Bern: Hans Huber.

Hicks, B. M., & Patrick, C. J. (2006). Psychopathy and Negative Emotionality: Analyses of Suppressor Effects Reveal Distinct Relations With Emotional Distress, Fearfulness, and Anger-Hostility. *Journal of Abnormal Psychology*, *115*(2), 276–287. doi:10.1037/0021-843X.115.2.276

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. doi:10.1007/BF02289447

Hu, X., Erb, M., Ackermann, H., Martin, J. A., Grodd, W., & Reiterer, S. M. (2011). Voxel-based morphometry studies of personality: issue of statistical model specification--effect of nuisance covariates. *NeuroImage*, *54*(3), 1994–2005. doi:10.1016/j.neuroimage.2010.10.024

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. *Biology of Mood & Anxiety Disorders*, *3*(1), 12. doi:10.1186/2045-5380-3-12

Jimura, K., Chushak, M. S., & Braver, T. S. (2013). Impulsivity and self-control during intertemporal decision making linked to the neural dynamics of reward value representation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(1), 344–357. doi:10.1523/JNEUROSCI.0919-12.2013

Joseph, J. E., Liu, X., Jiang, Y., Lynam, D., & Kelly, T. H. (2009). Neural correlates of emotional reactivity in sensation seeking. *Psychological Science*, *20*(2), 215–223. doi:10.1111/j.1467-9280.2009.02283.x

Jupp, B., & Dalley, J. W. (2013). Behavioral endophenotypes of drug addiction: Etiological insights from neuroimaging studies. *Neuropharmacology*. doi:10.1016/j.neuropharm.2013.05.041

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151. doi:10.1177/001316446002000116

Karrass, J., Walden, T. A., Conture, E. G., Graham, C. G., Arnold, H. S., Hartfield, K. N., & Schwenk, K. A. (2006). Relation of emotional reactivity and regulation to

childhood stuttering. *Journal of Communication Disorders*, *39*(6), 402–423. doi:10.1016/j.jcomdis.2005.12.004

Koelsch, S., Skouras, S., & Jentschke, S. (2013). Neural Correlates of Emotional Personality: A Structural and Functional Magnetic Resonance Imaging Study. *PLoS ONE*, *8*(11), e77196. doi:10.1371/journal.pone.0077196

Lauriola, M., & Levin, I. P. (2001). Personality traits and risky decision-making in a controlled experimental task: an exploratory study. *Personality and Individual Differences*, *31*(2), 215–226. doi:10.1016/S0191-8869(00)00130-6

Lauriola, M., Panno, A., Levin, I. P., & Lejuez, C. W. (2013). Individual Differences in Risky Decision Making: A Meta-analysis of Sensation Seeking and Impulsivity with the Balloon Analogue Risk Task. *Journal of Behavioral Decision Making*, n/a–n/a. doi:10.1002/bdm.1784

Lauriola, M., Panno, A., Levin, I. P., & Lejuez, C. W. (2014). Individual Differences in Risky Decision Making: A Meta-analysis of Sensation Seeking and Impulsivity with the Balloon Analogue Risk Task. *Journal of Behavioral Decision Making*, *27*(1), 20–36. doi:10.1002/bdm.1784

Lee, D. (2013). Decision making: from neuroscience to psychiatry. *Neuron*, *78*(2), 233–248. doi:10.1016/j.neuron.2013.04.008

Lim, S.-L., O'Doherty, J. P., & Rangel, A. (2011). The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That is Guided by Visual Attention. *The Journal of Neuroscience*, *31*(37), 13214–13223. doi:10.1523/JNEUROSCI.1246-11.2011

Matsuo, K., Nicoletti, M., Nemoto, K., Hatch, J. P., Peluso, M. A. M., Nery, F. G., & Soares, J. C. (2009). A voxel-based morphometry study of frontal gray matter correlates of impulsivity. *Human Brain Mapping*, *30*(4), 1188–1195. doi:10.1002/hbm.20588

Morishima, Y., Schunk, D., Bruhin, A., Ruff, C. C., & Fehr, E. (2012). Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron*, *75*(1), 73–79. doi:10.1016/j.neuron.2012.05.021

Munafò, M. R., & Flint, J. (2011). Dissecting the genetic architecture of human personality. *Trends in Cognitive Sciences*, *15*(9), 395–400. doi:10.1016/j.tics.2011.07.007

Pallanti, S., & Bernardi, S. (2009). Neurobiology of repeated transcranial magnetic stimulation in the treatment of anxiety: a critical review. *International Clinical Psychopharmacology*, *24*(4), 163–173. doi:10.1097/YIC.0b013e32832c2639

Pascual-Leone, A., Rubio, B., Pallardó, F., & Catalá, M. D. (1996). Rapid-rate transcranial magnetic stimulation of left dorsolateral prefrontal cortex in drug-resistant depression. *The Lancet*, *348*(9022), 233–237. doi:10.1016/S0140-6736(96)01219-6

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*(6), 768–774.

Paulus, M. P., & Stein, M. B. (2006). An Insular View of Anxiety. *Biological Psychiatry*, *60*(4), 383–387. doi:10.1016/j.biopsych.2006.03.042

Paulus, M. P., & Yu, A. J. (2012). Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, *16*(9), 476–483. doi:10.1016/j.tics.2012.07.009

Pine, A., Shiner, T., Seymour, B., & Dolan, R. J. (2010). Dopamine, time, and impulsivity in humans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(26), 8888–8896. doi:10.1523/JNEUROSCI.6028-09.2010

Pitman, R. K., Shin, L. M., & Rauch, S. L. (2001). Investigating the pathogenesis of posttraumatic stress disorder with neuroimaging. *The Journal of Clinical Psychiatry*, *62 Suppl 17*, 47–54.

Reddy, L. F., Lee, J., Davis, M. C., Altshuler, L., Glahn, D. C., Miklowitz, D. J., & Green, M. F. (2014). Impulsivity and Risk Taking in Bipolar Disorder and Schizophrenia. *Neuropsychopharmacology*, *39*(2), 456–463. doi:10.1038/npp.2013.218

Robbins, T. W., Gillan, C. M., Smith, D. G., de Wit, S., & Ersche, K. D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends in Cognitive Sciences*, *16*(1), 81–91. doi:10.1016/j.tics.2011.11.009

Rosenberg, P. B., Mehndiratta, R. B., Mehndiratta, Y. P., Wamer, A., Rosse, R. B., & Balish, M. (2002). Repetitive transcranial magnetic stimulation treatment of

comorbid posttraumatic stress disorder and major depression. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *14*(3), 270–276.

Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, *70*(6), 1054–1069. doi:10.1016/j.neuron.2011.05.014

Schilling, C., Kühn, S., Romanowski, A., Schubert, F., Kathmann, N., & Gallinat, J. (2012). Cortical thickness correlates with impulsiveness in healthy adults. *NeuroImage*, *59*(1), 824–830. doi:10.1016/j.neuroimage.2011.07.058

Schutter, D. J. L. G., & van Honk, J. (2005). A framework for targeting alternative brain regions with repetitive transcranial magnetic stimulation in the treatment of depression. *Journal of Psychiatry & Neuroscience: JPN*, *30*(2), 91–97.

Schutter, D. J. L. G., Van Honk, J., Koppeschaar, H., & Kahn, R. (2002). Cortisol and reduced interhemispheric coupling between the left prefrontal and the right parietal cortex. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *14*(1), 89–90.

Spampinato, M. D., Wood, P. D., De Simone, M. D., & Grafman, P. D. (2009). Neural Correlates of Anxiety in Healthy Volunteers: A Voxel-Based Morphometry Study. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *21*(2), 199–205.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Sripada, C. S., Gonzalez, R., Phan, K. L., & Liberzon, I. (2011). The neural correlates of intertemporal decision-making: contributions of subjective value, stimulus type, and trait impulsivity. *Human Brain Mapping*, *32*(10), 1637–1648. doi:10.1002/hbm.21136

Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., & Woolard, J. (2008). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: Evidence for a dual systems model. *Developmental Psychology*, *44*(6), 1764–1778. doi:10.1037/a0012955

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a

macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289. doi:10.1006/nimg.2001.0978

Van Schuerbeek, P., Baeken, C., De Raedt, R., De Mey, J., & Luypaert, R. (2011). Individual differences in local gray and white matter volumes reflect differences in temperament and character: A voxel-based morphometry study in healthy young females. *Brain Research*, *1371*, 32–42. doi:10.1016/j.brainres.2010.11.073

Vasic, N., Walter, H., Sambataro, F., & Wolf, R. C. (2009). Aberrant functional connectivity of dorsolateral prefrontal and cingulate networks in patients with major depression during working memory processing. *Psychological Medicine*, *39*(6), 977–987. doi:10.1017/S0033291708004443

Wagner, G., Koch, K., Schachtzabel, C., Reichenbach, J. R., Sauer, H., & Schlösser Md, R. G. M. (2008). Enhanced rostral anterior cingulate cortex activation during cognitive control is related to orbitofrontal volume reduction in unipolar depression. *Journal of Psychiatry & Neuroscience: JPN*, *33*(3), 199–208.

Walton, M. E., Behrens, T. E. J., Noonan, M. P., & Rushworth, M. F. S. (2011). Giving credit where credit is due: orbitofrontal cortex and valuation in an uncertain world. *Annals of the New York Academy of Sciences*, *1239*, 14–24. doi:10.1111/j.1749-6632.2011.06257.x

Weber, E. U., Blais, A.-R., & Betz, N. E. (2002). *A Domain-Specific Risk-Attitude Scale: Measuring Risk Perceptions and Risk Behaviors* (SSRN Scholarly Paper No. ID 1422152). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=1422152

Whiteside, S. P., & Lynam, D. R. (2001). The Five Factor Model and impulsivity: using a structural model of personality to understand impulsivity. *Personality and Individual Differences*, *30*(4), 669–689. doi:10.1016/S0191-8869(00)00064-7

Wilbertz, G., van Elst, L. T., Delgado, M. R., Maier, S., Feige, B., Philipsen, A., & Blechert, J. (2012). Orbitofrontal reward sensitivity and impulsivity in adult attention deficit hyperactivity disorder. *NeuroImage*, *60*(1), 353–361. doi:10.1016/j.neuroimage.2011.12.011

Wittmann, M., & Paulus, M. P. (2008). Decision making, impulsivity and time perception. *Trends in Cognitive Sciences*, *12*(1), 7–12. doi:10.1016/j.tics.2007.10.004

Wolf, R. C., Thomann, P. A., Sambataro, F., Vasic, N., Schmid, M., & Wolf, N. D. (2012). Orbitofrontal cortex and impulsivity in borderline personality disorder: an MRI study of baseline brain perfusion. *European Archives of Psychiatry and Clinical Neuroscience*, *262*(8), 677–685. doi:10.1007/s00406-012-0303-1

Wolkenstein, L., & Plewnia, C. (2013). Amelioration of Cognitive Control in Depression by Transcranial Direct Current Stimulation. *Biological Psychiatry*, *73*(7), 646–651. doi:10.1016/j.biopsych.2012.10.010

Yamasue, H., Abe, O., Suga, M., Yamada, H., Inoue, H., Tochigi, M., … Kasai, K. (2008). Gender-Common and -Specific Neuroanatomical Basis of Human Anxiety-Related Personality Traits. *Cerebral Cortex*, *18*(1), 46–52. doi:10.1093/cercor/bhm030

Zuckerman, M. (1996a). Item revisions in the Sensation Seeking Scale Form V (SSS-V). *Personality and Individual Differences*, *20*(4), 515. doi:10.1016/0191-8869(95)00195-6

Zuckerman, M. (1996b). The psychobiological model for impulsive unsocialized sensation seeking: a comparative approach. *Neuropsychobiology*, *34*(3), 125–129.

Zuckerman, M., Eysenck, S. B., & J, H. (1978). Sensation seeking in England and America: Cross-cultural, age, and sex comparisons. *Journal of Consulting and Clinical Psychology*, *46*(1), 139–149. doi:10.1037/0022-006X.46.1.139

List of tables and figures

Tables

Figures